Physiological changes induced by gene overexpression in plant cell cultures: characterization and statistical significance

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades einer

Doktorin der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Lea Anna Isgard Vaas

aus Stade

1. Referentin oder Referent:	PD Dr. Hans Peter Klenk
2. Referentin oder Referent:	Professor Dr. Ralf-Rainer Mendel
eingereicht am:	21. 12. 2011
mündliche Prüfung (Disputation) am:	31. 05. 2012

Druckjahr 2012

Zusammenfassung

Im Bereich der Pflanzenbiologie haben Werkzeuge der funktionalen Genomik das Verständnis von Stress-Signal-Wahrnehmung und -Weiterleitung vertieft und Einblicke in die beteiligten molekularen Netzwerke, besonders bei abiotischen Stressoren wie Trockenheit oder Salz-Stress, ermöglicht. Heterotrophe, dedifferentierte Suspensions-Zellkulturen erlauben es Stressoren homogen auf die Zellen wirken zu lassen und eine Charakterisierung unabhängig von morphologischen Strukturen vorzunehmen. Im Zusammenhang mit Pathogenität untersuchte Proteine (Pathogenesis-related (PR) proteins), besonders der Untergruppe 10a sind bekannt dafür sowohl unter Einfluss von biotischem als auch abiotischem Stress mit erhöhter Expression zu reagieren. Obwohl dieses Phänomen in verschiedenen Pflanzenarten, -organen, -gewebetypen und Entwicklungsstadien beobachtet werden konnte, ist die eigentliche Funktion des Proteins sowie seine Interaktionspartner weitgehend unerforscht.

In Forschungsvorhaben mit transgenen Pflanzen gibt es drei Arbeitsschritte: (i) Genetisch modifizierte Organismen (Linien) werden etabliert, (ii) vergleichende Experimente werden angestellt um das differenzierende Verhalten der genetisch modifizierten Organismen zu charakterisieren und (iii) es werden individuell zugeschnittene Analyse-Strategien und Software genutzt um die gewonnen Daten auszuwerten. Basierend auf einem Set von heterotrophen, dedifferenzierten Suspensions-Zellkulturen von *Solanum tuberosum* cv. Désirée (Wild-typ und *pr-10a* überexprimierende transgene Linien), deckt diese Arbeit alle drei Schritte ab und erweitert darüber hinausgehend das bisher angewandte Methodenspektrum in allen diesen Arbeitsschritten. Es werden neben der Vektor-Architektur, Kryokonservierung als neues Experimentelles Design zur Erhebung von neuartigen physiologischen Informationen erörtert und eine adäquate Software-Lösung zur einfach anzuwendenden Visualisierung und Daten-Analyse vorgestellt.

Besonders das Konzept von Respirationsmessungen stellt eine wertvolle Ergänzung des bisher vorhandenen Methodenspektrums zur Untersuchung von Phänotypen in heterotrophen dedifferentierten Zellkulturen dar. Darauf aufbauend steht die Veröffentlichung eines frei erhältliches R-Pakets bevor, um die Auswertung von Respirationskinetiken mit dem hier vorgestellten Methodenspektrum in eine umfassende Software-Lösung zu integrieren. Auf diese Weise wird eine tiefgehende Analyse, die signifikant mehr biologisch relevante Informationen liefert als die Auswertung von Endpunktmessungen oder dichotomen Daten, einem breiten Publikum ermöglicht.

Abstract

In plant-cell-systems, functional-genomics tools advanced our understanding of stress-signal perception and transduction and exhibited unrivalled insights in the associated molecular regulatory network especially regarding abiotic stress such as drought and salt stress. Heterotrophic, dedifferentiated suspended cell cultures facilitates the control of stress homogeneity and the characterisation of the cells' behaviour independently of plant morphologies. Pathogenesis-related (PR) proteins of group 10 and especially their member PR-10a were reported to show induced gene expression or protein abundance caused by several biotic and abiotic stressors in diverse studies on a variety of plants, organs, tissues and developmental stages. But the mode of action of the PR-10a protein itself as well as the pathways it could interfere with are still less well explored

Research using transgenic plants usually involves at least three steps: (i) the establishment of the genetically modified organism (GMO); (ii) applying experimental frameworks for analysing the behaviour of the GMO, usually in comparison to unmodified cells; and (iii) the application of software tailored to analysing the gathered data. Based on a set of heterotrophic dedifferentiated cell cultures of *Solanum tuberosum* cv. Désirée (wild-type and transgenic lines homologously overexpressing *pr-10a*) providing the main biological system to be analysed, this thesis covers all these steps not only by applying them, but also by augmenting the previously established approaches with methodological improvements. The the handling of the vector architecture, cryopreservation as a novel experimental framework for gathering information on GMO, the development of a method for collecting additional physiological data, and last but not least the elaboration of suitable software solutions for convenient data visualization and analysis are investigated.

Especially, the concept of respiration measurements is shown to be a meaningful complement for the set of methods when examining phenotypes of heterotrophic dedifferentiated cell cultures. The novel method provides valuable information in situations not targeted by more common methods such as growth measurements and the investigated data analyses strategies form the basis for a forthcoming freely available R package for the analysis of PM data, representative for respiration measurements resulting in longitudinal data sets. The availability of convenient and robust data exploration techniques via freely available software such as R will allow users to conduct in-depth data analyses that go significantly beyond the consideration of mere endpoint measurements and presence/absence calls.

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

El-Banna, A., Hajirezaei, M.R., Wissing, J., Ali, Z., Vaas, L., Heine-Dobbernack, E., Jacobsen, H.J., Schumacher, H.M., Kiesecker, H. Over-expression of PR-10a leads to increased salt and osmotic tolerance in potato cell cultures. Journal of Biotechnology 150: 277-287 (2010)

Tagungsbeiträge

Heine-Dobbernack, E., Seufert, S., Schumacher, H.-M., Vaas, L.A.I.: Preserving the physiological behaviour of potato cell cultures by freezing (Poster). COST Action – 871 Fundamental aspects of plant cryopreservation/ cryoprotection and genetic stability, 4th meeting of working group 1, Poznan (2010)

Heine-Dobbernack, E., Vaas, L.A.I. and Schumacher, H.-M.: Technical problems evaluating cryopreservation success in basic research (Vortrag). SLTB – Symposium, London (2011)

Schumacher, H.-M., Heine-Dobbernack, E. and Vaas, L.A.I.: Do we need cryopreservation in basic research projects using dedifferentiated plant cell cultures? (Vortrag) COST Action – 871 Fundamental aspects of plant cryopreservation/cryoprotection and genetic stability, 4th meeting of working group 1, Poznan (2010)

Vaas, L.A.I. and Göker, M.: Phenotype Microarray – New technology, new realm of high dimensional data. (Vortrag) Statistical Computing; Meeting of the working groups “Statistical Computing” and “Biostatistics” of the International Biometric Society (German Region), Günzburg (2010)

Vaas, L.A.I., Heine-Dobbernack, E., El-Banna, A., Kiesecker, H., Schumacher, H.-M.: Transgene expression monitoring in plant cell cultures (Poster). Botanikertagung, Berlin (2011)

Vaas, L.A.I., Marheine, M., Heine-Dobbernack, E., Schumacher, H.-M., Kiesecker, H.: Power and limitations of dicistronic vector systems for plant transformations (Vortrag). 7th

International Symposium on In Vitro Culture and Horticultural Breeding (IVCHB2011), Ghent (2011)

Vaas, L.A.I., Sikorski, J. and Göker, M.: Phenotype Microarray – Data organisation and analysis of respiration curves. (Vortrag) Statistical Computing; Meeting of the working groups “Statistical Computing” and “Biostatistics” of the International Biometric Society (German Region), Günzburg (2011)

Table of contents

Publications originating from this thesis

This thesis had originally been scheduled as a cumulative doctoral project. During the last three years, results from my work were used for writing a total of four manuscripts, three of which already were submitted for publication to international peer-reviewed scientific journals. All of them received overall positive responses from the reviewers and are now in the status “accepted for publication with minor revisions” as listed below:

- Vaas LAI, Sikorski J, Michael V, Göker M, Klenk H-P (20xx) Exploiting the wealth of Phenotype MicroArray data: efficacious visualization of and robust parameter estimation from respiration kinetics. (PLoS ONE manuscript no. PONE-D-11-15572R1, accepted for publication with minor revisions; confirmation e-mail on December 9th 2011)

- Vaas LAI, Marheine M, Seufert S, Schumacher H-M, Kiesecker H, Heine-Dobbernack E (20xx) Impact of *pr-10a* overexpression on the cryopreservation success of *Solanum tuberosum* suspension cultures. (Plant Cell Reports manuscript no. PCR-Sep-11-0621-M.R1, accepted for publications with minor revisions; confirmation e-mail on November 18th, 2011)

- Vaas LAI, Marheine M, Heine-Dobbernack E, Schumacher H-M, Kiesecker H (20xx) Power and limitations of dicistronic vector systems for plant transformations. (Acta Horticulturae, accepted for publication with minor revisions; confirmation e-mail on November 20th, 2011)

My fourth manuscript as first author, “Impacts of *pr-10a* overexpression at the molecular and the phenotypic level” is in preparation for submission to “International Journal of Molecular Biology”. The correspondence between these manuscripts and the chapters of this thesis is explained in detail in the introduction section.

Due to delays in the reviewing process in combination with time pressure caused by the changing dissertation regulatory (the TU Braunschweig changed its regulations during my Ph.D. work), this thesis is submitted at this time point as a non-cumulative work. Nevertheless, the submitted manuscripts apparently all satisfied the requirements for publication in peer-reviewed international scientific journals.

I. Introduction

Higher plants, in contrast to other living beings such as most animals, are characterised by a sessile life-form. They are obliged to manage all the environmental effects they are faced with at the place they are residing, since they cannot simply move to another, possibly more favourable, location. Taking into account that the first land plants already occurred in the basal Ordovician (Gensel 2008), this life form obviously developed efficient stress-managing strategies long time ago and perfected and conserved them on the genomic level. The genome comprehends a blueprint for an organism, but actually it is nothing more than a list of features storing the organisms' programme for life. That is, information about the spatial or temporal coordination of molecular (inter-)actions that result in the intended phenotype, provided that they are encoded in the genome, could not yet be deciphered (Quackenbush 2007). Even for the human genome, among the eukaryotes probably the best investigated today, the prediction of phenotypes from genomic data is still error-prone and restricted to only a few distinct types of genetic changes such as single-nucleotide polymorphisms (Lewis et al. 2011).

In this respect, the recent publication of the complete genome sequence of *Solanum tuberosum* (Xu et al. 2011) augmented the set of available completed plant genomes (see http://www.plantgdb.org for an up-to-date overview) by just another list of organisms' features.

However, in plant sciences, especially in breeding research, the application of genetic engineering and the evaluation of the resulting transgenic plants under stress conditions provides a valuable tool to learn how plants organize stress responses and deal with environmental effects (Somerville and Somerville 1999). But breeding approaches aim primarily at the improvement of crops (e.g., increased yield) and not directly at gaining basic functional insights (Bhatnagar-Mathur 2008). Here, genetic engineering, also known as the functional genomic approach (Somerville and Somerville 1999), provides a unique tool kit for the investigation of plants' stress responses on the molecular, cellular, or even higher organisational levels (Vinocur and Altman 2005).

Nevertheless, for tackling the fundamental problem of how the genomic program is realized as the organism grows, adapts, and responds to a wide range of stimuli, traditional approaches from physiology are indispensable (Edmeades et al. 2004). Especially regarding abiotic stresses such as drought and salt stress, which cause major limitations to crop productivity (Munns 2002), these functional-genomics tools advanced our understanding of stress-signal

perception and transduction and exhibited unrivalled insights in the associated molecular regulatory network (Valliyodan and Nguyen 2006). The first and most difficult step is to comprise and understand mechanisms by which plants perceive and transmit stress signals before they can initiate suitable responses (Xiong and Zhu 2001).

Pathogenesis-related (PR) proteins of group 10 (Liu and Ekramoddoullah 2006) and especially their member PR-10a were reported to show induced gene expression or protein abundance caused by several biotic and abiotic stressors in diverse studies on a variety of plants (Jelloulli et al. 2008), organs (Constable and Brisson 1995), tissues (Utriainen et al. 1998; Hashimoto et al. 2004; Mur et al. 2004; Ukaji et al. 2004) and developmental stages (Crowell et al. 1992; Walter et al. 1996). But the mode of action of the PR-10a protein itself as well as the pathways it could interfere with are still less well explored (Liu and Ekramoddoullah 2006; Gonzales-Lamothe et al. 2008).

Because elevated amounts of PR-10a proteins are repeatedly found in salt and osmotically stressed entire plants as well as cell cultures (Kav et al. 2004; Jain et al. 2006; Jellouli et al. 2008; Jellouli et al. 2010), a beneficial role in stress perception or signal transduction has been postulated (Boyle and Brisson 2001). Since signal perception and transduction pathways form a highly complex network rather than distinct and straightforward pathways (Urano et al. 2010), the genetic modification of a component involved in signal transduction or perception usually results in unexpected and puzzling new phenotypes (Zhu 2002; Colditz et al. 2007). The PR-10a protein represents such a case, too, since modulating the expression of *pr-10a* by genetic engineering yielded inconsistent results (reviewed in Liu and Ekramoddoullah 2006). Whereas some studies reported enhanced salt and/or osmotic tolerance due to *pr-10a* overexpression (Srivastava et al. 2004; El-Banna et al. 2010), the results in the context of pathogen attack were not easily interpretable. In potato plants overexpressing *pr-10a* no increase in the resistance against either *Phytophthora infestans* or potato virus X (Constable et al. 1993) was observed, whereas in the legume *Medicago truncatula* silencing of PR-10-like proteins increased the tolerance against infection with *Aphanomyces euteiches* (Colditz et al. 2007).

When intact plants are facing salt and/or drought stress, they are able to decelerate or modify the impact of this condition by complex interactions between different organs and tissues (Gaspar et al. 2002; Munns and Tester 2008). These in large parts unknown and complex spatial and temporal interactions at the complete organism’s level define the stress responses

at the cellular level and thus make a precise exploration of cause-and-effect chains rather difficult (Urano et al. 2010).

One strategy to facilitate exploratory insight can be the simplification of the testing system (Gaspar et al. 2002). Heterotrophic, dedifferentiated suspended cell cultures afford a higher precision at the cellular level since no such superordinate mechanisms against salt and/drought stress are realized. In such a system, the cells are surrounded by the medium and thus all of them experience equal conditions at the same time. This facilitates the control of stress homogeneity and the characterisation of the cells' behaviour independent of plant morphologies and the above-mentioned complexities. For example, it is possible to distinguish between the effect of ionic and non-ionic solutes supplementing the media and to measure their adverse effects by comparatively simple methods such as determining dry-weight accumulation (growth) or viability using staining assays (Lutts et al. 2004). Moreover, genetic-engineering tools are quickly and easily applicable when no plant regeneration step has to be passed (Ali et al. 2010; El-Banna et al. 2010).

In this thesis, the stress response characteristics of heterotrophic, dedifferentiated wild-type cell cultures of *Solanum tuberosum* cv. Désirée and two transgenic lines homologously overexpressing *pr-10a* based on the work of El-Banna et al. (2010) provide the main biological system to be analysed. The thesis, however, has also a strong methodological focus, and *pr-10a* and related topics are used as exemplars for elucidating novel approaches to data collection analysis.

After acquainting with the features of the dicistronic vector system used for the transformation in the second chapter, the third chapter employs the non-habitual framework of a cryopreservation procedure for gaining insights of the cells' response to osmotic stress. The regrowth rate after freezing and thawing, together with differential *pr-10a* gene-expression data, offers a remarkable possibility to learn about the cells' stress-response strategies.

The fourth chapter broadens the insights into the cells' response to osmotic and salt stress using classical growth-challenge experiments together with gene-expression monitoring over time. Additionally, the need for the measurement of the phenotypic features that determine the cells' very first reaction to a stress is motivated, and via the modification of a cell-viability assay a method meeting the outlined requirements is introduced.

Finally, chapter five deals with analysis strategies for the now achievable respiration kinetics, which are due to their longitudinal nature increasingly complex but also encode additional information. Suitable software solutions for convenient data visualization and quantitative

statistical analysis, i.e. methods aiming at the extraction of as many valuable insights as possible from high-dimensional longitudinal data are explored and presented.

Research using transgenic plants usually involves at least three steps: (i) the establishment of the genetically modified organism (GMO); (ii) applying experimental frameworks for analysing the behaviour of the GMO, usually in comparison to unmodified cells; and (iii) the application of software tailored to analysing the gathered data. This thesis covers all these steps not only by applying them but also by augmenting the previously established approaches with methodological improvements. The thesis investigates the handling of the vector architecture (second chapter), cryopreservation as a novel experimental framework for gathering information on GMO (third chapter), the development of a method for collecting additional physiological data (fourth chapter), and last but not least the elaboration of suitable software solutions for convenient data visualization and analysis meeting the requirements raised in the previous chapter (chapter five).

Impact of vector architecture on reporter-gene activity and resulting performance of the selection system

In the second chapter of this thesis, corresponding to the manuscript entitled “Power and limitations of dicistronic vector systems for plant transformations”, which is peer-reviewed and accepted for publication with minor revisions in ACTA HORTICULTURAE, practically relevant effects of the vector architecture of dicistronic constructs on the reporter-gene activity were investigated.

Genetic engineering, i.e., the stable incorporation of a gene of interest (GOI) into the plant cells’ genome, requires suitable selection systems for the transformed cells. Usually, the GOI is linked to a marker gene providing a selective advantage (Walden et al. 1990). However, the ultimate goal is to verify the successful expression of the GOI by, e.g., PCR-based relative quantification of expression and Northern-blot analysis for the quantification of transcription. Unfortunately, this requires substantial amounts of cell material and is destructive, resulting in a laborious and tedious screening process. Dicistronic vector systems including a firefly luciferase as reporter gene (RG) have been suggested to simplify the screening (Ali et al. 2010), as they enable both the selection of transgenic cells and the non-destructive detection of cells actively expressing the GOI via the chemiluminescence of luciferase (Joersbo and Okkels 1996).

Although its molecular nature remained unclear, recent results (Ali et al. 2010) indicated a potential effect of the combination of the specific GOI with a specific IRES element on the efficiency of the RG. Interestingly, findings obtained with mammalian cells also indicate an impact of the 1st cistron sequence on the activity of the IRES element and, thus, on the resulting amount of the RG protein from the 2nd cistron (Hennecke et al. 2001; Xia and Holcik 2009). However, because applications of dicistronic constructs in plants were few, knowledge about the impact of the vector architecture and the resulting physiological characteristics affecting the host-cell system is limited. This work aims at establishing dicistronic vector systems, which have been mainly used in animals, also to the plant world. The first step in this work, therefore, is to examine to what extent (a) the nature of the IRES element, (b) the gene product of the GOI and (c) the nucleotide sequence of the GOI impacts the detection of a successful transformation of plant cells, i.e. the strength of the luciferase-induced chemiluminescence emission. These parameters are addressed at the levels of transcription, protein amount and protein activity.

Combining the attained insights, this is the first study in plants that indicates a major impact of the 1st cistron important for both the translation of the reporter gene and the observable luciferase activity, highly similar to phenomena of IRES regulation described in mammalian cells (Hennecke et al. 2001).

Cryopreservation as a unique framework to evaluate the differences in cell lines' responses to dehydration and freezing

Especially when working with transgenic cell lines, the conservation of this cell material is of applied interest, since in culture expression levels can be altered due to the mere age of the culture or gene-silencing processes (Menges and Murray 2004). Thus, the thesis' third chapter, corresponding to the manuscript entitled "Impact of *pr-10a* overexpression on the cryopreservation success of *Solanum tuberosum* suspension cultures" and accepted after peer-review for publication with minor revisions in Plant Cell Reports, demonstrates that the treatment during the cryopreservation procedure provides an interesting basic research framework for evaluation of stress responses in the cells.

A special problem concerning the cryopreservation of dedifferentiated cell cultures is their higher vacuolization and, hence, higher water content compared to cells from meristematic plant tissues (Kaczmarcyk et al. 2011). Therefore, a substantial dehydration of cells is essential for their successful cryopreservation (Chetverikova 2008). Unfortunately, increased dehydration results in more severe chemical and mechanical damages of the cells (Kaviani

2011). As such, cells with good osmotic and salt tolerance have a higher chance to survive the pretreatment phase (Volk 2010).

As homologous overexpression of *pr-10a* in a cell line of *Solanum tuberosum* cv. Désirée resulted in increased survival under salt and osmotic challenge compared to the wild-type cells (El-Banna et al. 2010), the cryopreservation procedure provides a unique framework to evaluate the differences in the cell lines' responses to dehydration and freezing.

Sorbitol- and dimethylsulfoxide- (DMSO) based pretreatments are modified in terms of sorbitol concentration and treatment duration in order to determine the resulting *pr-10a* expression patterns in both wild type and transgenic cells. The pretreatment parameters are correlated with *pr-10a* gene-expression levels and cryopreservation success. Further, the role of *pr-10a* as a biomarker for the prediction of the success in cryopreservation is discussed. This is the first study that explores the impact of pr10a overexpression in the applied context of response to cryopreservation stress.

Bridging the gap between molecular and phenotypic data in stress challenges using cell cultures

In the fourth chapter, corresponding to the manuscript entitled "Impacts of *pr-10a* overexpression at the molecular and the phenotypic level" and in preparation for submission, a detailed investigation of the cell lines' response to both osmotic and salt stress is depicted.

The potato wild-type cell line (*Solanum tuberosum* cv. Désirée) is compared to two cell cultures homologously overexpressing the *pr-10a* gene. A a detailed analysis regarding the relative gene-expression patterns of *pr-10a* as well as *sebf* and *pti4*, whose gene products are involved in the *pr-10a* repressosome, provides insights at the molecular level into the response of heterotrophic cells to osmotic and salt-stress conditions.

In addition to realization of the established, relatively slowly developing growth curves, measuring phenotypic features during short-term cell responses is of interest. Obviously, heterotrophic dedifferentiated cell cultures have a limited repertoire of phenotypes compared to whole plants. Therefore, a method for monitoring the cells' respiration behaviour using a modified cell-viability assay is adopted in this work. In addition to the common growth curves, this approach is shown to provide information about the very first phase of a stress response, which is unattainable by growth measurements.

Furthermore, this technique is likely to facilitate studies on the metabolic activities of cells under even severe stress conditions or other adverse circumstances (Lutts et al. 2004), since it

was revealed that in plant cells respiration can occur independently of cell growth (Patterson 1979; Chapin 1991). (For instance, for the data at hand, the technique reveals that the *pr-10a* overexpression at least for severe 0.32M NaCl stress is not exclusively beneficial and thus provides valuable information about the impact of artificially enhanced presence of *pr-10a* on cell physiology.)

At this point, considerations about data analysis already have to be emphasized, because with the improvement of measurement techniques the longitudinal character of the output data becomes more important. That is, data complexity increases, yielding additional information coded in the shape characteristics of the curves (Brisbin et al. 1987). These curve features can unravel fundamental differences or similarities in the respiration behaviours under distinct treatments, which may be identifiable only by multi-parameter curve comparisons rather than by the simple comparison of means of end points. This is the first study that explores the impact of *pr-10a* overexpression in the context of molecular and phenotypic responses to salt osmotic stress.

Coming to terms with longitudinal data characteristics and increasing outcome complexity

In chapter five, corresponding to the manuscript entitled “Exploiting the wealth of Phenotype MicroArray data: efficacious visualization of and robust parameter estimation from respiration kinetics”, peer-reviewed and accepted for publication with minor revisions in PLoS One, techniques dealing with increasing data complexity, as caused by the improvement of measurement techniques by adding a longitudinal character to the data, are explored. To be able to optimally exploit the intrinsic information, it must be extracted from the shapes of the recorded curves and appropriately displayed.

Basically, a major biological feature, the phenotype, was until recently not accessible with high-throughput techniques. This is unfortunate, as it is the phenotype which is the object of selection and, hence, is the level at which evolutionary directions are governed (Mayr 1997), whereas all previously addressed “-omics” techniques merely study components which finally contribute to the phenotype (Papin 2004).

The Phenotype MicroArray (PM) system appears to close the gap of capturing a large number of phenotypes in high-throughput-systems. In this approach, a physiological reaction leading to the production of NADH engenders a redox potential and the flow of electrons to reduce a tetrazolium dye (Bochner and Savageau 1977) such as tetrazolium violet, thereby producing

purple colour. The more rapid this metabolic flow, i.e., cellular respiration, the more rapidly purple colour is formed (Bochner et al. 2001; Bochner 2009). The OmniLog™ PM system records the colour change every 15 minutes in an automated setting under up to 2000 distinct physiological challenges, such as the metabolism of single carbon sources, metabolism under varying osmolyte concentrations, and response to varying growth inhibitory substances (Bochner et al. 2001; Bochner 2009).

In common -omics techniques the recorded value is a qualitative (at best quantitative) information on the difference between two experiments, usually obtained from measurements at a single time point, commonly the end point (Lay et al. 2006). In contrast, the PM respiration kinetics adds a longitudinal dimension. This higher level of data complexity contains additional valuable biological information coded in the shape characteristics of the recorded curves in analogy to conventional growth curves (Brisbin et al. 1987). These curve features can, in principle, unravel fundamental differences or similarities in the respiration behaviour of distinct organisms, which cannot be identified by (identical) end-point measurements alone.

Motivated by the arising need for a suitable software solution, the free software environment R (R Development Core Team 2010) is explored for both data visualization and fitting of growth curves and thus facilitating a comparative analysis of PM data. R is one of the most widely used solutions for statistical computing, featuring powerful interactive data exploration as well as programming tools and numerous add-on packages.

First, the suitability of the *lattice* package (Sarkar 2008) is assessed for (re-)implementing and comparing previously published (Jacobsen et al. 2007) and alternative strategies for raw data visualization of 9792 bacterial respiration curves. Second, it is examined which kinds of divergences from typical sigmoid growth curves occur, which kinds of artefacts might affect the reproducibility of the results and, hence, which basic data quality control measures are necessary and can be performed using the here presented software tools.

Third, following the model-fitting approach of (Fodor et al. 2005) the grofit package (Kahm et al. 2010) is assessed for automatically conducting model fits as well as model-free fits using spline smoothers. The robustness of both approaches when inferring curve parameters (and their confidence intervals) from PM data is compared with the current implementation in the native OmniLog™ PM analysis software (Biolog 2009) and the specific merits and deficiencies of either method are determined.

Fourth, the tools are applied to biological research questions relevant for establishing settings for OmniLog™ PM production runs, illustrating how the experimenter can detect significance and magnitude of differences between the considered curve parameters to ensure reproducibility of the results in accordance with predefined quality standards (Gardner and Altman 1986).

Finally, as another example for the post-processing of the inferred parameters, the curves are classified into characteristic shapes. In contrast to the typical dichotomization of PM curves into occurrence of respiration and lack thereof (Xue et al. 2011), here curve archetypes (Eugster and Leisch 2009) are inferred to explicitly address the question of how many, and which, classes of curve shapes optimally represent the data.

The R package “opm” that is currently developed at the DSMZ and motivated by the in this chapter presented PLoS One manuscript, is to my knowledge the best bioinformatic tool currently available for longitudinal respiration kinetic data. The manuscript for this R package is currently under preparation. Although the presented PLoS One manuscript deals primarily with bacterial respiration kinetics, the here explored graphical visualisation and statistical tools have successfully been applied to the respiration kinetics of the *Solanum tuberosum* cell lines presented in fourth chapter, bridging the gap between the third and the fourth chapter.

Literature

Ali Z, Schumacher HM, Heine-Dobbernack E, El-Banna A, Hafeez FY et al. (2010) Dicistronic binary vector system – A versatile tool for gene expression studies in cell cultures and plants. Journal of Biotechnology 145: 9-16

Bhatnagar-Mathur P, Vadez V, Sharma KK (2008) Transgenic improvement for abiotic stress tolerance in plants: retrospect and prospects. Plant Cell Reports 27: 411-424

BiOLOG Inc. (2009) Converter, file management software, parametric software, Phenotype MicroArray, user guide, part # 90333. Biolog Inc., Hayward

Bochner BR and Savageau MA (1977) Generalized indicator plate for genetic, metabolic, and taxonomic studies with microorganisms. Applied Environmental Microbiology 33: 434-444

Bochner BR, Gadzinski P, Panomitros E (2001) Phenotype MicroArrays for high throughput phenotypic testing and assay of gene function. Genome Research 11: 1246-1255

Bochner BR (2009) Global phenotypic characterization of bacteria. FEMS Microbiological Reviews 33: 191-205

Boyle B, Brisson N (2001) Repression of the defense gene PR-10a by the single-stranded DNA binding protein SEBF. The Plant Cell 13: 2525-2537

Brisbin IL, Collins CT, White GC, McCallum DA (1987) A new paradigm for the analysis and interpretation of growth data: the shape of things to come. The Auk 104: 434-444

Chapin FS (1991) Integrated responses of plants to stress. A centralized system of physiological responses. BioScience 40: 29-31

Chetverikova EP (2008) Dehydration in cryopreservation of moist plant tissues and seed maturation. Biophysics 53: 304-307

Colditz F, Niehaus K, Kajinski F (2007) Silencing of PR-10-like proteins in *Medicago truncatula* results in an antagonistic induction of other PR proteins an in an increased tolerance upon infection with the oomycete *Aphanomyces euteiches*. Planta 226: 57-71

Constabel CP, Bertrand C, Brisson N (1993) Transgenic potato plants overexpressing the pathogenesis-related STH-2 gene show unaltered susceptibility to *Phytophthora infestans* and potato virus X. Plant Molecular Biology 22: 775-782

Constabel CP, Brisson N (1995) Stigma- and vascular-specific expression of the *pr-10a* gene of potato: a novel pattern of expression of a pathogenesis-related gene. Molecular Plant-Microbe Interactions 8: 104-1138

Crowell DN, John ME, Russel D, Amasona R (1992) Characterization of a stressinduced, developmentally regulated gene family from soybean. Plant Molecular Biology 18: 459-466

Edmeades GO, McMaster GS, White JW, Campos H (2004) Genomics and the physiologist: bridging the gap between genes and crop response. Field Crops Research 90: 5-18

El-Banna A, Hajirezaei MR, Wissing J, Ali Z, Vaas L et al. (2010) Over-expression of PR-10a leads to increased salt and osmotic tolerance in potato cell cultures. Journal of Biotechnology 150: 277-287

Eugster MJA and Leisch F (2009) From Spider-man to Hero – archetypal analysis in R. Journal of Statistical Software 30: 1-23

Fodor IK, Holtz-Morris AE, McCutchen-Maloney SL (2005) Growth curve models for the analysis of phenotype arrays for a systems biology overview of *Yersinia pestis*. Joint Statistical Meetings Minneapolis, MN, United States August 6, 2005 through August 10, 2005

Gardner MJ and Altman DG (1986) Confidence intervals rather than P values: estimation rather than hypothesis testing. British Medical Journal 292: 746-750

Gaspar T, Franck T, Bisbis B, Kevers C, Jouve L et al. (2002) Concepts in plant stress physiology. Application to plant tissue cultures. Plant Growth Regulation 37: 263-285

Gensel PG (2008) The earliest land plants. Annual Review of Ecology, Evolution, and Systematics 39: 459-477

Gonzales-Lamothe R, Boyle P, Dulude A, Roy V, Lezin-Doumbou C et al. (2008) The transcriptional activator Pti4 is required for the recruitment of a repressosome nucleated by repressor SEBF at the potato *pr-10*a gene. The Plant Cell 20: 3136-3147

Hashimoto M, Kisseleva L, Sawa S, Furukawa T, Komatsu S et al. (2004) A novel rice PR10 protein, RSOsPR10, specifically induced in roots by biotic and abiotic stresses, possibly via the jasmonic acid signaling pathway. Plant Cell Physiology 45: 550-559

Hennecke M, Kwissa M, Metzger K, Oumard A, Kröger A et al. (2001) Composition and arrangement of genes define the strength of IRES-driven translation in bicistronic mRNAs. Nucleic Acids Research 29: 3327-3334

Jacobsen JS, Joyner DC, Borglin SE, Hazen TC, Arkin AP et al. (2007) Visualization of growth curve data from phenotype microarray experiments.11th International Conference on Information Visualization (IV07), Zürich, Switzerland, July 4-6, 2007. Published by the IEEE Computer Society

Jain S, Srivastava S, Sarin NB, Kav NNV (2006) Proteomics reveals elevated levels of PR10 proteins in saline-tolerant peanut (*Arachis hypogaea*) calli. Plant Physiology and Biochemistry 44: 253-259

Jellouli N, Ben Jouira H, Skouri H, Ghorbel A, Gourgouri A et al. (2008) Proteomic analysis of Tunisian grapevine cultivar Razegui under salt stress. Journal of Plant Physiology 165: 471-481

Jellouli N, Jouira BH, Daldoul S, Chenennaoui S, Ghorbel A et al. (2010) Proteomic and transcriptomic analysis of grapevine PR10 expression during salt stress and functional characterization in yeast. Plant Molecular Biology Reporter 28: 1-8

Joersbo M, Okkels FT (1996) A novel principle for selection of transgenic plant cells: positive selection. Plant Cell Reports 16:219-221

Kaczmarcyk A, Rokka VM, Keller ERJ (2011) Potato shoot tip cryopreservation. A review. Potato Research 54: 45-79

Kahm M, Hasenbrink G, Lichtenberg-Frate H, Ludwig J, Kschischo M (2010). Grofit: fitting biological growth curves with R. Journal of Statistical Software 33: 1-21

Kav NNV, Srivastava S, Goonewardene L, Blade SF (2004) Proteome-level changes in the roots of *Pisum sativum* in response to salinity. Annals of Applied Biology 145: 217-230

Kaviani B (2011) Conservation of plant genetic resources by cryopreservation. Australian Journal of Crop Science 5: 778-800

Lay JO Jr, Borgmann S, Liyanage R, Wilkins CL (2006) Problems with the “omics”. Trends in Analytical Chemistry 25: 1046-1056

Lewis SN, Nsoesie E, Weeks C, Qiao D, Zhang L (2011) Prediction of Disease and Phenotype Associations from Genome-Wide Association Studies. PLoS ONE 6: e27175

Liu JJ, Ekramoddoullah AKM (2006) The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses. Physiological and Molecular Plant Pathology 68: 3-13

Lutts S, Almansouri M, Kinet JM (2004) Salinity and water stress have contrasting effects on the relationship between growth and cell viability during and after stress exposure in durum wheat callus. Plant Science 167: 9-18

Mayr E (1997) The objects of selection. Proceedings of the National Academy of Science USA 94: 2091-2094

Menges M and Murray JAH (2004) Cryopreservation of transformed and wild-type *Arabidopsis* and tobacco cell suspension cultures. The Plant Journal 37: 635-644

Munns R (2002) Comparative physiology of salt and water stress. Plant, Cell & Environment 25: 239-250

Munns R, Tester M (2008) Mechanisms of salinity tolerance. Annual Review of Plant Biology 59: 651-681

Mur LAJ, Sturgess FJ, Farrell GG, Draper J (2004) The AoPR10 promoter and certain endogenous PR10 genes respond to oxidative signals in *Arabidopsis*. Molecular Plant Pathology 5: 435-451

Papin JA, Reed JL, Palsson BO (2004) Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. TRENDS in Biochemical Sciences 29: 641-647

Patterson MK (1979) Measurements of growth and viability of cells in culture. Methods in Enzymology 58: 141-152

Quackenbush J (2007) Extracting biology from high-dimensional biological data. The Journal of Experimental Biology 210: 1507-1517

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna:

Sarkar D (2008) Lattice: Multivariate Data Visualization with R. Springer, New York

Somerville C, Somerville S (1999) Plant Functional Genomics. Science 285: 380-383

Srivastava S, Fristensky B, Kav NNV (2004) Constitutive expression of a PR10 protein enhances the germination of *Brassica napus* under saline conditions. Plant Cell Physiology 45: 1320-1324

Ukaji N, Kuwabara C, Takezawa D, Arakawa K, Fujikawa S (2004) Accumulation of pathogenesis-related (PR) 10/Bet v 1 protein homologues in mulberry (*Morus bombycis* Koidz.) during winter. Plant, Cell and Environment 27: 1112-1121

Urano K, Kurihara Y, Seki M, Shinozaki K (2010) ‘Omics’ analyses of regulatory networks in plant abiotic stress responses. Current Opinion in Plant Biology 13: 132-138

Utriainen M, Kokko H, Auriola S, Sarrazin O, Kärenlampi S (1998) PR-10 protein is induced by copper stress in roots and leaves of a Cu/Zn tolerant clone of birch, *Betula pendula*. Plant, Cell and Environment 21: 821-828

Valliyodan B, Nguyen HT (2006) Understanding regulatory networks and engineering for enhanced drought tolerance in plants. Current Opinion in Plant Biology 9: 189-195

Vinocur B, Altman A (2005) Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. Current Opinion in Biotechnology 16: 123-132

Volk GM (2010) Application of functional genomics and proteomics to plant cryopreservation. Current Genomics 11: 24-29

Walden R, Koncz C, Schell J (1990) The use of gene vectors in plant molecular biology. Methods Mol Cell Biol 1:175-194

Walter MH, Liu JW, Wünn J, Hess D (1996) Bean ribonuclease-like pathogenesis-related protein gene (Ypr10) display complex patterns of developmental, dark-induced and exogenous-stimulus-dependent expression. European Journal of Biochemistry 239: 281-293

Xia X, Holcik M (2009) Strong Eukaryotic IRESs have weak secondary structure. PLoS One 4:e4136

Xiong L, Zhu JK (2001) Abiotic stress signal transduction in plants: Molecular and genetic perspectives. Physiologia Plantarum 112: 152-166

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S et al. (2011) Genome sequence and analysis of the tuber crop potato. Nature 475: 189-195

Xue X, Sztajer H, Buddruhs N, Petersen J, Rohde M et al. (2011) Lack of the delta subunit of RNA polymerase increases virulence related traits of *Streptococcus mutans*. PLoS ONE 6: e20075

Zhu JK (2002) Salt and drought stress signal transduction in plants. Annual Review in Plant Biology 53: 247-273

II. Power and limitations of dicistronic vector systems for plant transformations

Abstract

Dicistronic vector constructs harboring an internal ribosomal entry site (IRES) element facilitate the coordinated co-expression of physically independent proteins, as, e.g., a target protein along with a reporter protein. Dicistronic vector constructs have been studied intensively in mammalian cell systems, but their efficiency and applicability in plant cell systems is much less known. The influence of the 1st cistron nucleotide sequence on the translational expression level of the reporter gene encoded by the 2nd cistron was investigated. The testing set consists of the plant viral TMV IRES element and the mammalian viral polio IRES element in combination with either the *Atnhx* gene, coding for a Na^+/H^+ antiporter, or the *gus* gene. Using the resulting four combinations the impact on a firefly luciferase (*luc*) as reporter gene was studied at the transcriptional, translational and phenotypical level. In order to discriminate between nucleotide sequence of the 1st cistron and its gene product, we introduced a frame shift in the 1st cistron coding sequence. The sequence of the IRES element strongly influenced the translational reporter gene expression present in the 2nd cistron. The results show that the dicistronic vector construct has to be carefully designed and tested for each individual application to obtain the desired level of translational expression of the reporter gene.

Introduction

Genetic engineering, i.e., the stable incorporation of genes of interest (GOI) into the plant cells' genome, requires suitable selection systems for the transformed cells. Usually, the GOI is linked to a marker gene providing a selective advantage (Walden et al. 1990). However, the ultimate goal is to also verify the successful expression of the GOI by, e.g., PCR-based relative quantification of expression and northern blot analysis for the quantification of transcription. Unfortunately, this requires substantial amounts of cell material and is destructive, resulting in a laborious and tedious screening process. Dicistronic vectors systems including a firefly luciferase as reporter gene (RG) have recently been suggested to simplify the screening (Ali et al. 2010), as they enable both the selection of transgenic cells and the

non-destructive detection of cells actively expressing the GOI via the chemiluminescence of luciferase (Joersbo and Okkels 1996).

In dicistronic vector systems the GOI is located in a 1st cistron, physically linked with the RG in the 2nd cistron by an internal ribosomal entry site (IRES). While the translation of most eukaryotic mRNAs is initiated in the conventional cap-dependent way, it can alternatively be initiated using the binding of ribosomes to specific RNA structures, the IRESs (reviewed in Jackson 2005). In mammalian systems such constructs have been further elaborated (Dirks et al. 1994), and several distinct interactions of IRES elements with host cells and vector configurations have been explored (Kazadi et al. 2008). However, the transferability of the dicistronic vector concept into plant cells has only recently been addressed (El-Banna et al. 2010).

In the study of Ali et al. (2010) the At*nhx* gene (Na^+/H^+ antiporter) as GOI and the *luc* gene as RG were utilized to test the efficiencies of two distinct IRES elements, the IRES from the plant pathogen crucifer tobamo virus (TMV) and the mammalian *Poliovirus* IRES. The *gus* gene was used as GOI in combination with the TMV IRES element as proof of principle (Ali et al. 2010). While the At*nhx* gene in combination with the TMV IRES element yielded high luciferase activity in transformed tobacco cells, neither the combination of At*nhx* with the polio IRES element nor the combination of *gus* with the TMV IRES element resulted in significant activity (Ali et al. 2010).

Although its molecular nature remained unclear, the results indicated a potential effect of the combination of the specific GOI with a specific IRES element on the efficiency of the RG. Since physiological effects of the used target genes could not be excluded, the expressed protein of the GOI presumably interfered with the metabolism of the transformed cells to either increase (higher salt tolerance due to the At*nhx1* expression) or decrease (β-glucuronidase proteins in the cells) the signal strength of the RG (Ali et al. 2010).

Findings obtained with mammalian cells indicate an impact of the 1st cistron sequence on the activity of the IRES element which is reflected by the RG protein levels derived from the 2nd cistron (Hennecke et al. 2001; Xia and Holcik 2009).

In the present study the molecular nature of RG efficiency in dicistronic vector systems with tobacco plant cells transformed by leaf infiltration assays is analysed. A set of eight combinations of GOI with IRES elements was constructed to investigate the influence of the IRES elements, the vector architecture and the physiological effect of the target gene in the

plant cell system. Two 1st cistron gene sequences, the Na^+/H^+ antiporter gene At*nhx1* (AF056190) from *Arabidopsis thaliana*, and the *gus* gene (AY292368) from *Escherichia coli* in combination with a firefly luciferase (*luc*; from the basic pGII0229MAS luc vector, see Ali et al. 2010 for details) as RG in the 2nd cistron, were studied. To discriminate between the effects of the 1st cistron nucleotide sequence and the physiological effect of its gene product, a frame shift was introduced in the 1st codon of the target gene sequence to prevent the formation of physiologically active gene products. Two IRES elements, the TMV cp 148 IRES element from the crucifer tobamo virus (TMV), and the polio IRES element from the *Poliovirus* (polio), differing in their general activity (Dorokhov et al. 2002) were compared using the two 1st cistron genes along with their frame shift variants.

The dicistronic vector combinations were analysed by means of (a) luciferase activity in transiently transformed cells to verify the proper functioning, (b) quantification of transcription efficiency via (q)RT-PCR to exclude the possibility that different RG activities result from altered transcription rates, and (c) the translation efficiency of the GOI via Western blots to determine the amount of actually translated luciferase reporter protein.

Materials and Methods

Constructs and vector functionality

Based on the vector system introduced in Ali et al. (2010), either the At*nhx1* (AF056190) or the *gus* gene as GOI, were combined with either the TMV cp148 or the polio IRES element (Figure 1). To introduce a frame shift in the 1st codon of the At*nhx1* gene the vector pGII0229MAS AtNhx IRES luc was digested with the restriction enzyme *Pst*1. Protruding ends were blunted with DNA blunting enzyme (CloneJET™ PCR Cloning Kit, Fermentas), introducing a frame shift. The vector was re-ligated to yield pGII0229MAS AtNhxF IRES luc. The frame shift in the 1st codon of the *gus* sequence was obtained via amplification of the *gus* sequence from the vector using a forward primer with sequence modification, introducing an *Nco*1 restriction site along with the frame shift: modified sequence ncoI (f) 5'-ATGCCGGGCCATGGTTACGTCCTGTAGAAACCCC-3', gus mscI (r) 5'-CGCATCACGCAGTTCAACGCTGACATCACC-3'. The modified amplification product was cloned with *Sma*1 and *Msc*1 into the original vector to yield the new vector pGII0229MAS gusF IRES luc. The successful introduction of the frame shift was verified by sequencing the respective region.

Vector functionality was tested via a leaf infiltration assay according to English et al. (1997) as described in Ali et al. (2010). As negative control for the Western blot analysis leaves transformed with vectors not containing any insert were processed.

Semiquantitative luciferase measurement

To qualitatively monitor the transient *luc* expression, transgenic leaves were washed with 0.5% Tween 20 in water solution, rinsed with distilled water and dried on a filter paper. D-luciferin potassium salt (Cat. Nr. bc219, Synchem, Felsberg, Germany) was dissolved in water (1mM) and sprayed on the leaves. The luminescence was recorded with a Fuji LAS 3000 imager and processed with the Aida® quantification software. Leaf material was stored in liquid nitrogen for further analysis.

RNA isolation and quantitative real-time PCR

Total RNA was extracted from leaf samples (100mg fresh weight) after grinding under liquid nitrogen using TriFast Gold (Peqlab, Erlangen, Germany) according to the manufacturer's instructions. cDNA was synthesised using the RevertAid 1st Strand Synthesis kit (Fermentas, St. Leon Roth, Germany). The mono- and di-cistronic mRNA was quantified by PCR TaqMan of the *luc* gene (primer: luc141(f) 5'-TATGAACATTTCGCAGCCTA-3', luc141(r) 5'-ATCGACTGAAATCCCTGGTA-3', probe: luc141 5'-TGTTTCCAAAAAGGGGTTGCAAA-3') was amplified. As internal standard primers against the *18S* rRNA gene 18S138(f) 5'-TAAAGGAATTGACGGAAGGG-3', 18S138 (r) 5'-CACCACCACCCATAGAATCA-3' probe: 18S138 5'-CGCAGGCTCCACTCCTGGTG-3' were used. Quantitative real-time PCR was performed on an Eppendorf Mastercycler ep realplex^4 platform by the following program: 30 s 95°C, [40× (95 °C, 5s.; 60°C, 20s.)]

Calculation of dCT values was done using spreadsheets. Values were calculated for luc standardized to the *18S* threshold cycle (ΔCt) (Pfaffl 2001).

Protein isolation and Western blot

For protein extraction, 100mg of frozen grinded leave tissue were suspended in 800µl extraction buffer (0.7M sucrose, 0.5M Tris, 30mM HCl, 50mM EDTA, 0.1M KCl and 2% [v/v] 2-mercaptoethanol) (Hurkman and Tanaka 1986). The protein pellet was re-dissolved in 5M urea in 50mM DTT buffer. After centrifugation (10,000g, 10min), protein concentration was determined using 2D Quant kit (GE Healthcare, Germany) by mixing 50µg of clear supernatant with loading buffer (0.025M Tris-HCl, 4% SDS and 20% glycerol in a 1:1 ratio

and incubating at 95°C for 10min. The proteins were separated in 12% SDS-polyacrylamide gels and blotted onto a PVDF membrane (Millipore Corporation, Bedford, MA, USA). For immunodetection, the membrane was blocked with 4% skimmed milk in PBST buffer (0.14M NaCl, 8.1mM Na_2HPO_4, 2.7mM KCl, 1.47mM KH_2PO_4 and 0.05% TWEEN) over night followed by incubation with 1:500 diluted rabbit anti-luciferase polyclonal antibodies (MBL, Nagoya, Japan) in PBST for 2h. The membrane was washed 3 times with PBST buffer and incubated for 45min with ECL® Anti-rabbit IgG (horseradish peroxidase linked whole antibody, GE Healthcare UK) at a 1:2,000 dilution in PBST. The membrane was then washed with PBST for 45min and incubated in Western Lightning® ECL (Perkin Elmer). Visualization of the 61kDa luciferase band was performed by chemiluminescence detection using FUJI LAS 3000.

Results

Distinct levels of light emission in semi-quantitative luciferase assay

Irrespective of whether functional or disabled by a frameshift, the At*nhx1* gene in combination with the TMV cp148 IRES displayed a similarly high luciferase activity (Figure 2). In comparison to leaves transfected with the At*nhx1* constructs, leaves transfected with the gus-cp148 IRES and gusF-cp148 IRES constructs showed low luminescence levels, independent from the presence of the introduced frame shift in the 1st cistron gene. This indicates that the transcribed gene product of the 1st cistron gene does not influence the expression of the 2nd cistron.

Leaves transformed with constructs containing the polio IRES generally emitted much less chemiluminescence.

Gene transcription rates in quantitative real-time PCR

To ensure that the differences in luciferase activity (Figure 2) were not caused by distinct amounts of vector transcription, the level of transcription of the luciferase RG was analysed via Q-PCR. Figure 3 gives the amounts of luc-transcripts normalized against *18S* rRNA transcripts. Interestingly, no differences in transcript levels were detected for either construct.

Differences in actual protein yield detected in Western blots

To verify that the distinct levels of light emission (Figure 2) were caused by different amounts of protein, the yield of luciferase protein in the cells was quantified by immunodetection (Figure 4). The extracts from leaves transformed with the constructs harbouring the At*nhx1* gene followed by TMV cp148 IRES revealed a sizeable protein amount, regardless of the introduced frame shift. In contrast, neither samples from leaves transformed with pGII0229MAS gus TMVcp148 luc nor with pGII0229MAS gusF TMVcp148 luc exhibited detectable amounts of luciferase protein. Also, none of the samples from transformations using constructs harbouring the polio IRES displayed a detectable amount of luciferase protein.

Discussion

The objective of this study was to examine to what extent (a) the nature of the IRES element, (b) the gene product of the GOI and (c) the nucleotide sequence of the GOI impacts the detection of a successful transformation of plant cells, i.e. the strength of the luciferase-induced chemiluminescence emission. These parameters were assessed by determining the levels of transcription, translation and enzyme activity.

All constructs were successfully tested concerning their performance in transient expression assays and yielded equally large amounts of transcription products (Figure 3), indicating that neither the nature of the 1st cistron gene nor the type of IRES element had any measurable influence on the amount of transgenic mRNA molecules.

In contrast, we observed substantial differences at the level of translation, as only the combination of At*nhx1* gene and TMV IRES element yielded significant chemiluminescence and measurable amounts of luciferase protein, irrespective of the presence of a frame shift. All other combinations yielded neither apparent amounts of chemiluminescence nor luciferase protein (Figures. 2, 4).

No indication was found that the activity of the protein resulting from the 1st cistron gene translation alters the overall plant cell metabolism in a way affecting the chemiluminescence detection reaction. The 1st cistron At*nhx1* gene provides a Na^+/H^+ antiporter which may relieve salt stress by actively exporting Na^+ ions into either the extracellular space or into vacuoles (Blumwald and Poole 1987). In leaf infiltration assays, local salt stress is most likely

induced by the infiltration medium. However, the high luciferase activity was not dependent on the protein product derived from the 1st cistronic gene as the frame shift had no effect.

A potentially unfavorable physiological effect of the *gus* gene due to its glucuronidase function could not be confirmed, as the *gus*-containing constructs with a frame shift did not impede the putative adverse effect yielding in measurable chemiluminescence or amounts of protein.

Hence, the nature of the specific combination of GOI and IRES element may hinder successful translation, since the individual components alone, the IRES element and the *gus* gene, were successfully and efficiently applied in separate vector systems (Walden et al. 1990; Jackson et al. 2005), but, when combined, are only transcribed but not translated. This raises the question of how, on a molecular level, translation is inhibited. Xia and Holcik (2009) showed for eukaryotic IRES elements that strong activity is associated with weak secondary structure in the RNA molecule. We explored the strength of the secondary structure of the At*nhx* and *gus* gene mRNA by calculating their minimum folding energy (MFE) (using the Vienna RNAfold web server; Zuker and Stiegler 1981). Interestingly, we found a larger MFE for the *gus* gene in the 5' UTR upstream of the IRES initiation codon, suggesting a more stable secondary structure of the mRNA in this region (data not shown). This may result in steric hindrance of an optimal translation initiation by the IRES element attached to the ribosome and might provide an explanation for the observed results.

Until now it is unclear, which parameters determine the efficiency of either viral or cellular IRESs, nor are any common characteristics of these parameters known (Xia and Holcik 2009). Moreover, a general lack of observable IRES sequence similarity has resulted in the widely held believe that IRESs likely possess stable secondary structures, allowing them to interact with the components of the translation machinery (Xia and Holcik 2009). Other results suggested that the impact of the general trans-acting factors (eIFs) or IRES-specific trans-acting factors (ITAFs) could be the main control mechanism of IRES activity (Spriggs et al. 2005).

Thus, recently a concept was proposed in which viral IRESs were classified according to their structural diversity. Filbin and Kieft (2009) suggested that IRES RNAs with the most inherent stably folded structure are those that require the fewest factors. The less stable IRESs are, the more ITAFs and eIFs are needed. However, currently the literature disregards vector architecture, especially the characteristics of 1st and 2nd cistron genes.

In our study, the identity of the 1st cistron had a major impact and seems to be important for both the translation of the reporter gene and the observed luciferase activity similar to IRES regulation found in mammalian cells (Henneke et al. 2001). In our set-up, the utilized TMV IRES element originated from a virus which naturally can infect *Nicotiana* and thus fits perfectly to cells in the testing system using leaf infiltration. In contrast, the investigated IRES element from *Poliovirus*, which normally infects mammalian cells, may thus not exhibit strong influences of the 1st cistron gene sequence, since its overall activity is too low in tobacco cells. However, IRES elements can strongly affect the physiological state of mammalian host cells. Hence, cell type-dependent translational activity as well as host specificity may play a role for viral IRES elements (Honda et al. 2000; Qin and Sarnow 2004).

Although dicistronic vectors provide a simple monitoring system suitable for high-throughput analysis of correlations between transgenic expression and physiological traits, several issues have to be considered to optimize their usage.

The choice of a suitable IRES element is vitally important, as the specific combination of GOI and IRES element may fail to work, even though both elements alone are functional. We therefore recommend to test every construct in the target cell system and to carefully take into account the physiological status of the host cells when monitoring for transgenic cells under distinct conditions. When the GOI negatively effects the translational efficiency of the IRES element a strategy for improvement could be the insertion of a spacer element between the 1st cistron gene and the IRES element, which relaxes the strength of the secondary structure. In our system, an initial approach could be the introduction of a short sequence between the *gus* gene and the TMV IRES element to relax the hypothesized strength of the secondary structure of the mRNA. Vectors amended by only a short fragment would be not much longer than the initial one in contrast to approaches, where several IRES element-RG sequences are attached (Bouabe et al. 2008).

Literature

Ali Z, Schumacher HM, Heine-Dobbernack E, El-Banna A, Hafeez FY et al. (2010) Dicistronic binary vector system – A versatile tool for gene expression studies in cell cultures and plants. Journal of Biotechnology 145: 9-16

Blumwald E, Poole RJ (1987) Salt tolerance in suspension cultures of sugar beet: induction of Na/H antiport activity at the tonoplast by growth in salt. Plant Physiology 83: 884-887

Bouabe H, Fässler R, Heesemann J (2008) Improvement of reporter activity by IRES-mediated polycistronic reporter system. Nucleic Acids Research 36: e28

Dirks W, Schaper F, Kirchhoff S, Morelle C, Hauser H (1994) A multifunctional vector family for gene expression in mammalian cells. Gene 149: 387-388

Dorokhov YL, Skulachev MV, Ianov PA, Zvereva SD, Tjuklina LG, et al. (2002) Polypurine (A)-rich sequences promote cross-kingdom conservation of internal ribosome entry. Proceedings of the National Academy of Science USA 99: 5301-5306

El-Banna A, Hajirezaei MR, Wissing J, Zahid A, Vaas L et al. (2010) Over-expression of PR-10a leads to increased salt and osmotic tolerance in potato cell cultures. Journal of Biotechnology 150: 277-287

English JJ, Davenport GF, Elmayan T, Vaucheret H, Baulcombe DC (1997) Requirement of sense transcription for homology-dependent virus resistance and trans-inactivation. Plant Journal 12: 597-603

Filbin ME and Kieft JS (2009) Toward a structural understanding of IRES RNA function. Current Opinion in Structural Biology 19: 267-276

Henneke M, Kwissa M, Metzger K, Oumard A, Kröger A et al. (2001) Composition and arrangement of genes define the strength of IRES-driven translation in bicistronic mRNAs. Nucleic Acids Research 29: 3327-3334

Honda M, Kaneko S, Matsushita E, Kobayashi K, Abell GA et al. (2000) Cell cycle regulation of hepatitis C virus internal ribosomal entry site-directed translation. Gastroenterology 118: 152-162

Hurkman WJ and Tanaka CK (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. Plant Physiology 81: 802-806

Jackson RJ (2005) Alternative mechanisms of initiating translation of mammalian mRNAs. Biochemical Society Transaction 33: 1231-1241

Joersbo M and Okkels FT (1996) A novel principle for selection of transgenic plant cells: positive selection. Plant Cell Reports 16: 219-221

Kazadi K, Loeuillet C, Deutsch S, Siuffi A, Munoz M et al. (2008) Genomic determinants of the efficiency of internal ribosomal entry sites of viral and cellular origin. Nucleic Acids Research 36: 6918-6925

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Research 29: e45

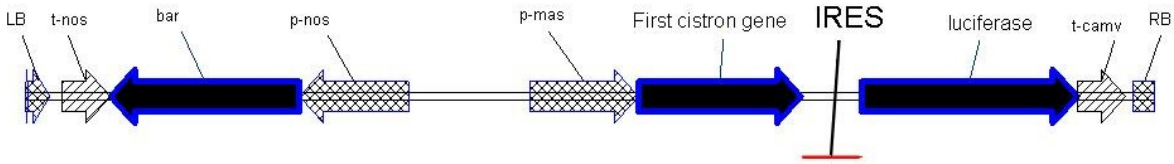
Qin X and Sarnow P (2004) Preferential translation of internal ribosome entry site-containing mRNAs during the mitotic cycle in mammalian cells. Jounal of Biological Chemistry 279: 13721-13728

Spriggs KA, Bushell M, Mitchell SA, Willis AE (2005) Internal ribosome entry segment-mediated translation during apoptosis: the role of IRES-trans-acting factors. Cell Death and Differentiation 12: 585-591

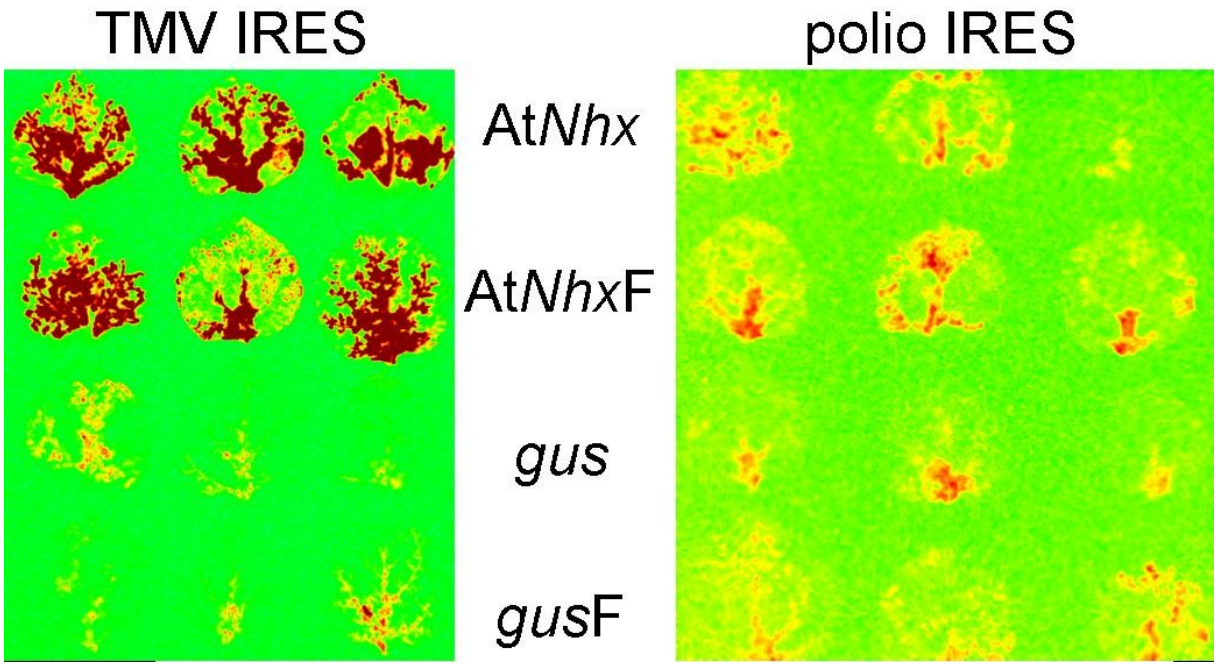
Walden R, Koncz C, Schell J (1990) The use of gene vectors in plant molecular biology. Methods in Molecular Cell Biology 1: 175-194

Xia X, Holcik M (2009) Strong Eukaryotic IRESs have weak secondary structure. PLoS One 4: e4136

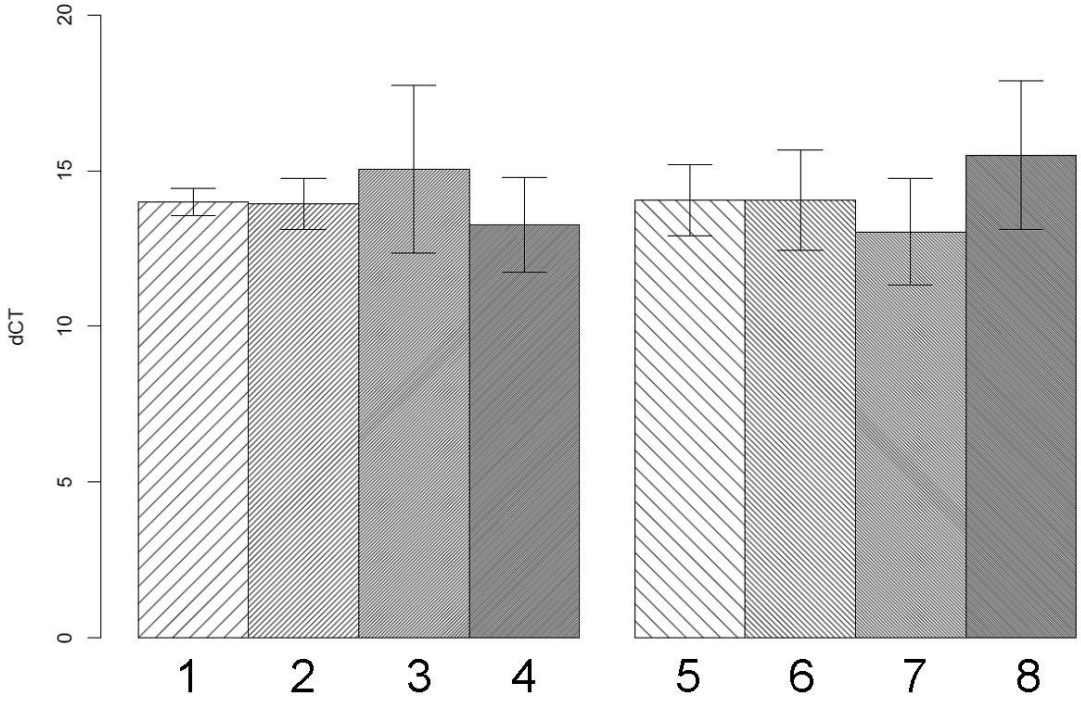
Zuker M, Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Research 9: 133-148



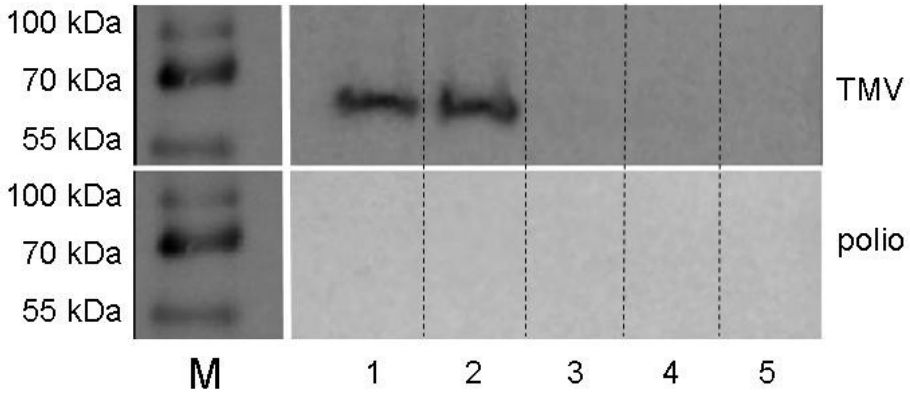
Chapter II, figure 1. Vector T-DNA cassette schema of the dicistronic vector pGII0229MAS 1st cistron IRES luc. For further details see Ali et al. (2010).



Chapter II, figure 2. Imaging of luciferase activities within the tobacco leaves. The colour shifts represent the increase in light intensity from low (yellow) to high (red).



Chapter II, figure 3. Relative quantification by Q-RT-PCR of luc expression normalized against *18S* rRNA; (1) Atnhx1 TMV luc (2) Atnhx1F TMV luc (3) gusTMV luc (4) gusF TMV luc (5) Atnhx1 polio luc (6) Atnhx1F polio luc (7) gus polio luc (8) gusF polio luc. The mean values for 3 individual samples are given. Error bars represent standard deviations.



Chapter II, figure 4. Western-blot analysis of luciferase proteins transiently expressed from dicistronic transcripts. Upper row: TMV IRES constructs, lower row: polio IRES constructs; M: Marker (1) Atnhx1 luc (2) Atnhx1F IRES luc (3) gus IRES luc (4) gusF IRES luc (5) control.

III. Impact of *pr-10a* overexpression on the cryopreservation success of *Solanum tuberosum* suspension cultures

Abstract

Although many genes are supposed to be a part of plant cell tolerance mechanisms against osmotic or salt stress, their influence on tolerance towards stress during cryopreservation procedures has rarely been investigated. For instance, the overexpression of the pathogenesis-related gene 10a (*pr-10a*) leads to improved osmotic tolerance in a transgenic cell culture of *Solanum tuberosum* cv. Désirée. In this study a cryopreservation method, consisting of osmotic pretreatment, cryoprotection with DMSO and controlled-rate freezing, was used to characterize the relation between cryopreservation success and *pr-10a* expression in suspension cultures of *Solanum tuberosum* wild-type cells and cells overexpressing pathogenesis-related protein 10a (Pr-10a).

By varying the sorbitol concentration, thus modifying the strength of the osmotic stress during the pretreatment phase, it can be shown that the wild type can successfully be cryopreserved only in a relatively narrow range of sorbitol concentrations, while the *pr-10a* overexpression leads to an enhanced cryopreservation success over the whole range of applied sorbitol concentrations. Together with transcription data we show that *pr-10a* overexpression causes an enhanced osmotic tolerance, which in turn leads to enhanced cryopreservability, but also indicates a role of *pr-10a* in signal transduction. An increased cryopreservability of the transgenic cell line occurs for pretreatments longer than 24 hours. Since both genotypes, characterized by distinct baseline levels of expression, exhibited similar patterns of expression induction, the induction of *pr-10a* appears to be a key step in the stress signal transduction of plant cells under osmotic stress.

Introduction

Cryopreservation is the method of choice for long-term storage of plant material (Benson 2008; Engelmann 2009). At a low annual cost the individual plants' geno- and phenotypic characteristics are securely preserved (Engelmann 2004; Kaviani 2011). Especially the conservation of transgenic cell material is of interest, since in culture expression levels can be altered due to the mere age of the culture or gene-silencing processes (Menges and Murray 2004).

Despite the promise of cryopreservation, effects such as cell damage caused by ice-crystal formation during freezing and/or thawing may attenuate success (Reed 2008; Benson 2008). Thus, a favourable cryopreservation method has to mitigate ice crystal formation by the removal of water or the addition of cryoprotectants (Fuller 2004; Reed 2008). A special problem concerning the cryopreservation of dedifferentiated cell cultures is their higher vacuolization and, hence, higher water content compared to cells from meristematic plant tissues (Kaczmarcyk et al. 2011). Therefore, a substantial dehydration of cells is essential for their successful cryopreservation (Chetverikova 2008). This is usually achieved by the vitrification of plant cell material using PVS2 solution (Sakai et al. 1990), the treatment with cryoprotectants such as Dimethylsulfoxide (DMSO) (Chen et al. 1984a) or a mixture of sucrose, glycerol and DMSO (DGS), as well as the addition of proline (Withers and King 1979) or carbohydrates such as sorbitol (Chen 1984b), trehalose (Bhandal et al. 1985) or sucrose (Lu et al. 2009). Unfortunately, increased dehydration results in more severe chemical and mechanical damages of the cells (Kaviani 2011). As such, cells with good osmotic and salt tolerance have a higher chance to survive the pretreatment phase (Volk 2010).

Figure 1 schematically illustrates the principal relationship between the beneficial and adverse effects of cell dehydration on cryopreservation success. It is highest under conditions where dehydration is strong enough to prevent ice crystal formation but still gentle enough to avoid strong direct cell damage.

The actual optimum depends on a variety of additional parameters such as the geographical and ecological origin of the plant material (Berjak et al. 2011), the genotype of the cell material (Schäfer-Menuhr and Schumacher 1997), and its age and physiological status (Harding 2009), and thus can be much more complex than illustrated above (Figure 1). The strength of the pretreatment is a combination of the concentration of the applied substance(s) and their residence time. The physiological changes during pretreatment were identified as a conglomerate of competences in cold adaptation, including the maintenance of membrane stability and the scavenging of free radicals after thawing (Harding 2009; Volk 2010), severely affecting cryopreservability.

In the past cryopreservation procedures were mainly developed based on empirical grounds and thus methods strongly depended on tissue and plant species/variety. With projects such as CRYMCEPT (EU 5th framework programme, Project Reference: QLK5-CT-2002-01279), the determination of physico- and biochemical state of the considered plant species/tissues needed for successful cryopreservation were elaborated in some detail (see Panis and Lambardi 2006

for an overview). With increasing information from fundamental research on causal relationships determining the cryopreservability of plant materials (Carpentier et al. 2007; Swennen et al 2011), awareness increased that physiological factors strongly determine the success of survival (Harding 2004; Harding 2010).

Here, we study the effect of the pretreatment on the success of cryopreservation by controlled rate freezing. To evaluate the physiological changes resulting from pretreatment quality on the success of cryopreservation, we compare wild type cells of *Solanum tuberosum* cv. Désirée with a transgenic cell line overexpressing the pathogenesis-related protein 10a (Pr-10a).

Pr-10a was primarily studied in the context of attacks by pathogens (Liu and Ekramoddoullah 2006) but is also known to play a role in higher stress tolerance induced by increased ABA levels or abiotic factors such as drought or salt stress (Salekdeh et al. 2002; Kav et al.; 2004; Srivastava et al. 2004; Jain et al. 2006; Liu and Ekramoddoullah 2006; Bahramnejad et al. 2010). Thus, the cryopreservation procedure provides a unique framework to evaluate the differences in the cell lines' responses to dehydration and freezing. We recently showed that homologous overexpression of *pr-10a* in a cell line of *Solanum tuberosum* cv. Désirée resulted in increased survival under salt and osmotic challenge compared to the wild type cells (El Banna et al. 2010). The concentrations of proline and oxidized glutathione, regarded as indicators for salt and/or osmotic tolerance (Ashraf and Harris 2004), were increased (El Banna et al. 2010). In the present study we modify sorbitol- and DMSO-based pretreatments in terms of sorbitol concentration and treatment duration and determine the resulting *pr-10a* expression patterns in both wild type and transgenic cells. We correlate pretreatment parameters, *pr-10a* gene expression levels and cryopreservation success. Further, the role of *pr-10a* as a biomarker for the prediction of the success in cryopreservation will be discussed.

Material and Methods

Plant material and cultivation

The wild type of *Solanum tuberosum* cv. Désirée (WT, DSMZ No. PC-1182) and a transgenic line homologously overexpressing *pr-10a* (GMO) were subcultured and harvested as previously described (El Banna et al. 2010). For all experiments, suspension cells were taken from the logarithmic growth phase, 3d after the last subculturing.

Cryopreservation

Cryopreservation was performed applying a classical controlled-rate cooling approach based on the method of Withers and King (1980), comprising (1) dehydration pretreatment of cells in medium supplemented with sorbitol, (2) cryoprotection by exposure to DMSO, (3) start of the freezing process by slow cooling, (4) rapid immersion of the samples in liquid nitrogen, (5) storage in liquid nitrogen, (6) rapid thawing, and (7) recovery growth. Figure 2 schematically summarizes this procedure.

These cryopreservation experiments were performed in a minitest system previously worked out at DSMZ (Heine-Dobbernack et al. 2008).

For dehydration pretreatment, a cell inoculum with a fresh weight of 0.41g was subjected to 0.9ml of either 4X medium (control) (after Gamborg et al. (1968) containing 2 mg l-1 2,4-dichlorophenoxyacetic acid, 0.5 mg l-1 indole-3-acetic acid 0.5 mg l-1 1-naphtylacetic acid and 0.4 mg l-1 kinetin, pH 5.6) or 4X medium supplemented with different sorbitol concentrations (0.3M-1.2M), corresponding to a packed cell volume (pcv) of 30.0% in 24er multiwell plates (Greiner, Frickenhausen, Germany). The plates were incubated at 23°C for 48h on a shaker at 300rpm (Heidolph Titramx 101, Kehlheim, Germany). For each sorbitol concentration, six replicates were performed with both wild-type and transgenic cells. For cryoprotection, the 24 multiwell plates were precooled for 22min in a cool pack in a 4°C cold room on a shaker (Heidolph Titramx 1000, Kehlheim, Germany; speed 3.5rpm). To the cells in each well (plates on ice), 69µl precooled DMSO (final concentration 5% v/v) was added. The plates were incubated on ice in a cold room at 4°C on the shaker for 50min. Afterwards the suspension of each well was filled into one cryotube, which was placed on ice until the total DMSO incubation time (1.5h) had passed. For slow cooling, the cryo-samples were placed into the cooling chamber of a Planer Kryo10 automatical freezer (Planer PLC, Middlesex, UK). Starting at 4°C, the cells were cooled with a rate of -0.25°C/min down to -40°C. After a holding time of 15min at -40°C the cryotubes were immediately transferred into liquid nitrogen (-196°C) in a cryocontainer.

Samples for the determination of gene expression were treated identically, but instead of filling the suspended cells into the cryotubes, the supernatant in the wells was sucked off using tamponades (1.2 cm diameter, cut in half) after 1.5h total incubation time of DMSO had passed. A total of twelve wells were prepared for each sorbitol concentration. To obtain enough cell mass, the cells of each three corresponding wells were pooled together. The analysis was then accomplished with four independent samples from the different sorbitol

treatments for both the wild-type and the transgenic cell line. After harvesting, the cells were packed in aluminium foil, immediately immersed in liquid nitrogen and stored at -80°C until further analysis. As control group, cells were exposed to the normal pregrowth treatment including the 5% DMSO treatment for 1.5h, but without any sorbitol supplement in the medium. Thus, the measured changes in cryopreservation and *pr-10a* expression, if any, could be attributed to the distinct sorbitol concentrations.

For the experiments using distinct times of optimal sorbitol concentration for dehydration pretreatment a fresh weight of 11.93g cells, corresponding to a packed cell volume (pcv) of 30.0% was transferred into round sterile polystyrol jars (6.5cm high and 6.5cm in diameter; Greiner, Frickenhausen, Germany). A volume of 26ml 4X medium supplemented with the optimal sorbitol concentration as determined earlier was added to each jar. The jars were sealed with Parafilm M (Brand, Wertheim, Germany) to minimize evaporation and incubated for different intervals from 0h to 72h on a gyratory shaker (TR-250, Infors AG, Basel, Switzerland) with 50mm orbit (100 rpm) at 23°C.

For cryoprotection, 2ml precooled DMSO (final concentration 5% v/v) was added to the cells which had been precooled to 3.6°C for 11min on a shaker (Infors HT, Bottmingen, Switzerland) equipped with a cooling device (Julabo GmbH, Seelbach, Germany). Each jar was incubated for 50min at 3.6°C and 140rpm. Afterwards the jar was placed in an ice bath and aliquots of 1.5ml cell suspension per incubation time were filled into 2ml cryotubes (Nunc, Roskilde, Denmark). Each cryotube was placed on ice immediately after filling and when 1.5h total incubation time of DMSO had passed proceeded to slow cooling as described above. The control group in this experiment were cells sampled right before the pretreatment started. The measured changes in *pr-10a* expression, if any, could thus be ascribed to the combined effect of both the sorbitol and the DMSO treatment at the given time.

Regrowth behaviour without freezing

Cells were treated as described above, but after 1.5h total incubation time of DMSO had passed the content of the cryotubes was poured on 4X agar covered with three sterile filter paper discs (Filter discs (quant.), diameter 45mm, grade 391; Sartorius Stedim Biotech S.A., Aubagne, France) respectively. After 2h the uppermost filter was transferred to fresh 4X agar, and the agar plates were incubated for 15d in the dark at 23°C. Recovery growth was measured by dry weight.

Rapid thawing

Regrowth behaviour after freezing was determined by quickly thawing the cryotubes at 40°C in a water bath after storage in liquid nitrogen for 1d. As soon as the last visible ice had melted, the content of the cryotubes was poured on 4X agar covered with 3 sterile filter paper discs (see above). After 2h the uppermost filter was transferred to fresh 4X agar and the agar plates were incubated for 15d in the dark at 23°C. Recovery growth was measured by dry weight.

Dry-weight determination

The quality of cryopreservation was evaluated by determining the dry weight of the re-grown cells. Cells grown on a single disc of filter paper (see above) were quantitatively transferred to a new, pre-washed (0.9% NaCl + 0.05% HCHO), dried (60°C for 3d), and pre-weighed filter. After drying to constant weight at 60°C for 3d, the resulting weight minus the pre-defined filter weight gave the net cell dry mass.

RNA Isolation and quantitative real-time PCR

Total RNA was extracted from samples (100 mg fresh weight) with TriFast Gold (Peqlab, Erlangen, Germany) according to the manufacturer's instructions. cDNA synthesis was carried out with the RevertAid First Strand Synthesis kit (Fermentas, St. Leon Roth, Germany). For relative quantification of the *pr-10a* and luciferase mRNA by PCR TaqMan probes against the *pr-10a* sequence (primer pr-10a129(f) 5'-TACACATGAAGCCACAAGCA-3', pr-10a129(r) 5'-ATGCTTCCATCTCCCTCAGT-3', probe: pr-10a129 5'-TCAAAGCTTTGGTTGTTGATGCTGA-3') were used.As internal standard primers against the *18S* rRNA gene 18S138 (f) 5'-TAAAGGAATTGACGGAAGGG-3', 18S138 (r) 5'-CACCACCACCCATAGAATCA-3', probe: 18S138 5'-CGCAGGCTCCACTCCTGGTG-3' were used according to Nicot et al. (2005). Quantitative real-time PCR was performed on an Eppendorf Mastercycler ep realplex4 platform using the following program: 30s 95°C; [40× (95 °C, 5s; 60°C, 20s)].

As control group in the experiment for evaluating the impact of the distinct sorbitol concentrations, cells were exposed to the normal pregrowth treatment including the 5% DMSO treatment for 1.5h, but without any sorbitol supplement in the medium. Thus the measured changes in expression, if any, could be attributed to the sorbitol concentration.

Since the aim of the second experiment, in which cells were treated with 0.6M sorbitol for different pretreatment intervals, was to investigate the expression changes compared to the normal level within the cells, the control group in this experiment were cells sampled right before the start of the pregrowth treatment. The measured changes in expression, if any, could thus be ascribed to the combined effect of the sorbitol treatment at the given time.

Calculation of both dCT and ddCT values was done using spreadsheets and the statistical software R (R Development Core Team 2010). Values were calculated as fold change $2^{-\Delta\Delta Ct}$ for *pr-10a* standardized to the *18S* threshold cycle (ΔCt) with the differences between the treated cells and cells from control group. A fold change of 1 indicates no change in *pr-10a* expression caused by treatment (Pfaffl 2001).

Statistical analysis

For comparisons of means simultaneous multiple comparison procedures according (Hothorn et al. 2008) and the methods described in Schaarschmidt and Vaas (2009) were performed using statistical software R (R Development Core Team 2010).

Results

Cryopreservation success depends on the sorbitol concentration

The effect of distinct sorbitol concentrations during a 48h long pretreatment phase on the quality of cryopreservability is shown in Figure 3.

Apparently, treatment with sorbitol was essential for a successful cryopreservation of both the WT and the GMO. The WT showed reasonable regrowth only under the 0.4M, 0.6M and 0.8M sorbitol concentrations (Figure 3). In contrast, the GMO showed reasonable regrowth at all applied sorbitol concentrations, with a maximum at 0.4M and 0.6M (Figure 3) and was always superior to the wild type cells, except the control group, where no regrowth was observed at all.

Distinct *pr-10a* expression in WT and GMO at distinct sorbitol concentrations

To investigate the impact of the different sorbitol concentrations on the *pr-10a* gene expression, *pr-10a* expressions were measured after 48h pre-treatment duration via quantitative real-time PCR. As control group cells were exposed to the normal 4X medium during the pregrowth time including the 1.5h long 5% DMSO supplementation.

The relative increase of *pr-10a* expression in both WT and GMO at distinct sorbitol concentrations compared to the control experiment with no sorbitol is shown in Figure 4a.

The low-level treatments of 0.3M and 0.4M sorbitol appear to decrease the *pr-10a* expression in both genotypes (Figure 4b). At 0.6M sorbitol only the WT showed a very slight induction (mean fold change ca. 1.3). At concentrations above 0.6M the WT reacted with explicit expression inductions, which increased with increasing sorbitol treatments. The *pr-10a* expression fold change of the GVO also rose at sorbitol treatments above 0.6M but to a much smaller extent than that of the wild type.

We further determined the absolute expression levels of *pr-10a* in WT and GMO at distinct sorbitol concentrations (Figure 4b). At nearly all sorbitol concentrations, and in the control assay without sorbitol, the absolute expression level of *pr-10a* was substantially higher in the GMO than in the WT. Only at 0.8M sorbitol the expression levels were similar in both genotypes (Figure 4 b).

Although the absolute level of expression was higher in the GMO (Figure 4b), its relative increase is much higher in the WT than in the GMO (Figure 4a). In addition to the measurable overexpression of *pr-10a* in the GMO, the sorbitol treatments of concentrations above 0.8M can further increase the expression of *pr-10a* (Figure 4b).

Effect of pretreatment duration on regrowth behaviour at the optimal sorbitol concentration without freezing

For both the WT and the GMO, 0.4M as well as 0.6M sorbitol treatments could have been regarded as the optimal concentration at a pretreatment duration of 48h (Figure 3a). But taking into account the gene expression data (Figure 4a and b) and a visual inspection of cell colour and consistency (Figure 7) in addition to the texture and the smell of cells (data not shown), the 0.6M sorbitol treatment was considered to be superior.

The next step in evaluation of pretreatment influence on the cell viabiliy was to keep the sorbitol concentration constant at 0.6M and to vary the pre-treatment duration. In order to examine, if cells are damaged by the pretreatment increasingly with increasing sorbitol treatment duration, the cryopreservation step was omitted and cells were directly plated for dry weight determination after the pre-treatment procedure.

At 0.6M sorbitol concentration there was no effect of pretreatment duration on the regrowth of the WT (Figure 5). The GMO, in contrast, is characterised by slightly better regrowth

capabilities, which increase, to some extent significantly, with increasing pre-treatment duration (Figure 5).

Effect of distinct durations of optimal sorbitol treatment on cryopreservation success

To further examine the impact of different pretreatment durations of the optimal concentration of 0.6M sorbitol, the cells' regrowth behaviour was determined.

With increasing pretreatment duration, the ability to survive and regrow after cryopreservation increased in both genotypes (Figure 6 a). For the WT cell line the optimal pretreatment time was between 24h and 48h (Figure 6D, central panel), whereas the GMO exhibited enhanced regrowth also for the 48h and 72h pretreatment durations (Figure 6d, lower panel). At short pretreatment times (4h to 12h) the accumulated dry weight of the WT did not differ significantly from the GMO's dry weight (Figure 6a and d, upper panels). As the CIs indicate (Figure 6 d, upper panel) above 24h pretreatment time the GMO cells' regrowth was superior to that from the WT. These results indicate that the *pr-10a* overexpression leads to better adaptation osmotic stress the longer the osmotic stress lasts (>24 hours), priming the cells for the imminent cryopreservation and regrowth process.

pr-10a gene expression at varying pretreatment durations but constant 0.6M (optimal) sorbitol concentration

To specify the effect of *pr-10a* in cryopreservation success, we further determined the relative expression and the course of the induction of the *pr-10a* transcription at varying pretreatment durations and constant 0.6M sorbitol concentration.

Interestingly, GMO cells responded with an enhanced *pr-10a* expression fold change to the 1.5h incubation in medium supplemented with 0.6M sorbitol plus 5% DMSO (ca. 16 times; Figure 6b), whereas the WT did not react during these first 1.5h of the treatment.

Furthermore, the GMO cells exhibited a larger and beyond 24h rather constant level of *pr-10a* expression fold change compared to the WT (Figure 6b). Examining the dCT values, the *pr-10a* expression increased with increasing time with a maximal induction between 12h and 24h in both genotypes (Figure 6c). In analogy to the results from above (Figure 4), the GMO exhibited a higher and more prolonged level of induction of expression (Figure 6c).

The GMO showed a greater response to the pretreatments resulting in more transcripts after even a short time of stress. Since the superiority in cryopreservation success takes places only

after at least 24h pretreatment, the source for enhanced regrowth is supposed to be a usual, but now somewhat more effective stress response of the cells.

Discussion

Intracellular ice-crystal formation is the most detrimental effect during deep freezing of living cells. Dehydration of the cells to a level at which vitrification of the protoplasts takes place is of particular importance for a successful cryopreservation (Fuller 2004; Reed 2008). Thus, cells are commonly exposed to severe dehydration stress in order to achieve survival during deep freezing (Chetverikova 2008). Hence, mechanisms increasing osmotic tolerance will allow cells to withstand a higher level of dehydration, resulting in better survival of deep freezing.

Starting with the here shown enhancement of cryopreservability due to *pr-10a* overexpression at the optimal pretreatments, we could also obtain improved cryopreservability after non-optimal pretreatments (Figure 3). While the enhanced cryopreservation success after pretreatment with the high sorbitol concentrations agrees well with findings of improved osmotic tolerances due to *pr-10a* overexpression (El Banna et al. 2010), reasonable regrowth was also accomplished after mild osmotic stress of 0.3M sorbitol, which can not be explained by the improved osmotic tolerance of the GMO. The *pr-10a* overexpression is thus supposed to enable the cells to survive the cryopreservation success as long as any osmotic pretreatment was applied. Note that the strength of this osmotic stress is to come second, since already 0.3M sorbitol pretreatment caused reasonable regrowth as well as the high concentrations did.

This rather supports the proposed role of Pr-10a in signal transduction (Liu and Ekramoddoullah 2006) initiating further stress response reactions in the cells. Although this involvement of PR10 proteins in stress responses has been reported very early (Matton and Brisson 1989), the evaluation of the expression behaviour over time and at different stress levels received only little attention (Liu and Ekramoddoullah 2006; Jelloulli et al. 2008; El-kereamy et al. 2009). Most efforts were devoted to elucidate regulation of gene expression and to characterize transcription factors (Després et al. 1995; Gonzalez-Lamothe et al. 2008). Interestingly, the shape of the curve relating the GMO's cryopreservation success to the different sorbitol concentrations was not changed, but the curve shifted upwards, indicating that the endogenous *pr-10a* gene was still normally responding to the applied treatments. Further, we could show that the *pr-10a* expression after 48h of treatment duration does not depend linearly on the stress level (Figure 4). Although the absolute level of expression is higher in the GMO (Figure 4b), its relative increase is much larger in the WT than in the

GMO (Figure 4a). Nevertheless, the GMO holds constantly higher amounts of measurable transcripts, which potentially lead to more Pr-10a protein triggering further stress responses even under the mild 0.3M sorbitol concentration.

Regarding the interaction between the optimal pretreatment with 0.6M sorbitol and its residence time, the WT's cryopreservation success benefited, surprisingly, only from the short pretreatment times up to 24h. In contrast, the GMO showed a high positive correlation between prolonged pretreatment time and cryopreservation success up to the maximum pretreatment residence time of 72h (Figure 6a). The gene expression analyses documented the genotypes' differences regarding the sensitivity against the pretreatment already after the first hour. Thus, the GMO is supposed to react faster and stronger to even short-term stress. The constitutive overexpression apparently provides increased amounts of *pr-10a* transcripts, but the endogenous *pr-10a* is still serviceable, and its induction caused the rise in *pr-10a* transcripts in analogy to the WT cells (Figure 4 and 6).

Interestingly, over the course of time we found a better coincidence of *pr-10a* induction and a measurable increase in cryopreservation success (Figure 6). While during the first 24h both genotypes increased the amount of *pr-10a* transcripts, both genotypes' cryopreservation success also increased. The WT's expression thereafter decreased again, and the cryopreservation success did not increase further, but the GMO held its *pr-10a* expression at the high level and increased its cryopreservation success further (Figure 6). Since no considerably adverse effect by the pure pretreatment could be found in the WT (Figure 5), the enhanced *pr-10a* expression is supposed to take its effect in the following steps of the cryopreservation procedure. A possible hormone mediating effect of the PR-10 proteins has been discussed earlier. A binding capacity of the Pr-10a protein to cytokinin (Fujimoto et al. 1998) or to brassinosteroids (Markovics Housley et al. 2003) has been demonstrated.

Beside the osmotic pressure, the applied substance itself may also act directly as signalling molecule in stress responses, as it has been reported for sorbitol-sugar interaction (Zhou et al. 2006) and also for sugar signalling (Gibson 2005; Gupta and Kaur 2005) and proline (Kishor et al. 2005). The resulting physiological changes in the cells would then be a conglomerate of answers to the increased osmotic pressure, the ionic disturbance in the cytosol due to dehydration, and the induced signalling pathways. Thus, we cannot exclude that the altered gene expressions were also influenced by sugar signalling and not primarily by the osmotic stress.

The elucidation of the induction of *pr-10a* gene expression patterns resulted in detailed information about the complex interactions between the treatment severities as well as the time course of the stress. Particularly monitoring the expression pattern over time exhibited that a certain amount of transcripts has to be available for a minimum period of time to ensure cryopreservation success. But the interactions between the induction of *pr-10a* gene expression and other signals triggered by the stress are too complex to enable one to determine a threshold for cryopreservation success or failure. This also supports the role of *pr-10a* in signal transduction (Liu and Ekramoddoullah 2006), initiating further stress response reactions in the cells.

When working with in-vitro cultures and especially with artificial systems such as cell cultures one has to be aware of physiological changes resulting from somaclonal variations. Undifferentiated cell cultures consist of a heterogeneous population of cells which probably always also contain genetically aberrant cells. But somaclonal variation events are most likely to occur during the transformation process (Filipecki and Malepszy 2006) rather than during maintenance under normal conditions.

The present study was performed with a transgenic cell line which showed consistently altered characteristics such as osmotic and salt tolerance as well as proline and glutathion metabolism (El Banna et al. 2010). With potato and peanut cell cultures it was shown that complicated and longlasting selection procedures can produce a comparable range in tolerances to osmotic and/or salt stresses, given that the selection procedure favours beneficial somaclonal variations (Sabbah and Tal 1990; Ochatt et al. 1999; Jain et al. 2006). But cell lines predominantly retain tolerance only when cultivated under continuous selection pressure (DSMZ, unpublished data). Since the cell lines considered in this study have never been exposed to salt or osmotic stress during the initiation, transformation or maintenance phase and retained tolerance also when cultivated on non-selective media, it is highly likely that the reason for the observed differences between wild type and transgenic cells is the *pr-10a* overexpression rather than somaclonal variation. Although the results from this study suggest an impact of the *pr-10a* overexpression on cryopreservation behavior, further studies are still necessary to clarify the molecular function of the Pr-10a protein and its interaction partners in vivo. This information would also help to understand whether molecular markers exist which could be used to predict success or failure of plant cell cryopreservation.

Literature

Ashraf M, Harris PJC (2004) Potential biochemical indicators of salinity tolerance in plants. Plant Science 166: 3-16

Bahramnejad B, Goodwin PH, Zhang J, Atnaseo C, Erickson LR (2010) A comparison of two class 10 pathogenesis-related genes from alfalfa and their activation by multiple stresses and stress-related signaling molecules. Plant Cell Reports 29: 1235-1250

Benson EE (2008) Cryopreservation of phytodiversity: A critical appraisal of theory and practise. Critical Reviews in Plant Sciences 27: 141-219

Berjak P, Bartels P. Benson EE, Harding K Mycock DJ et al. (2011) Cryoconservation of South African plant genetic diversity. In Vitro Cellular & Developmental Biology – Plant 47: 65-81

Bhandal IS, Hauptmann RM, Widholm JM (1985) Trehalose as cryoprotectant for the freeze preservation of carrot and tobacco cells. Plant Physiology 78: 430-432

Carpentier S, Witters E, Laukens K, Van Onckelen H, Swennen R et al. (2011) Banana (*Musa* spp.) as a model to study the meristem proteome: Acclimation to osmotic stress. Proteomics 7: 92-105

Chen THH, Kartha KK, Constabel F, Gusta LV (1984a) Freezing characteristics of cultured *Catharanthus roseus* (L.) G. Don cells treated with dimethylsulfoxide and sorbitol in relation to cryopreservation. Plant Physiology 75: 720-725

Chen THH, Kartha KK, Leung NL, Kurz WGW, Chatson KB et al. (1984b) Cryopreservation of alkaloid-producing cell cultures of periwinkle (*Catharanthus roseus*). Plant Physiology 75: 726-731

Chetverikova EP (2008) Dehydration in cryopreservation of moist plant tissues and seed maturation. Biophysics 53: 304-307

Després C, Subramaniam R, Matton DP, Brisson N (1995) The activation of the potato PR-lOa gene requires the phosphorylation of the nuclear factor PBFI. The Plant Cell 7: 589-598

El-Banna A, Hajirezaei MR, Wissing J, Ali Z, Vaas L et al. (2010) Over-expression of PR-10a leads to increased salt and osmotic tolerance in potato cell cultures. Journal of Biotechnology 150: 277-287

El-kereamy A, Jayasankar S, Taheri A, Errampalli D, Paliyath G (2009) Expression analysis of a plum pathogenesis related 10 (PR10) protein during brown rot infection. Plant Cell Reports 28: 95-102

Engelmann F (2009) Use of biotechnologies for conserving plant biodiversity. Acta Horticulturae 812: 63-81

Engelmann F (2004) Plant cryopreservation: Progress and prospects. In Vitro Cellular & Developmental Biology – Plant 40: 427-433

Filipecki M, Malepszy S (2006) Unintended consequences of plant transformation: a molecular insight. Journal of Applied Genetics 47: 277-286

Fujimoto Y, Nagata R, Fukasawa H, Yano K, Azumua M et al. (1998) Purification and cDNA cloning of cytokinin-specific binding protein from mung bean (*Vigna radiata*) European Journal of Biochemistry 258: 794-802

Fuller BJ (2004) Cryoprotectants: The essential antifreezes to protect life in the frozen state. CryoLetters 25: 375-388

Gamborg, OL, Miller RA, Ojima, K (1968) Nutrient requirements of suspension cultures of soy- bean root cells. Experimental Cell Research 50: 151-158

Gibson SI (2005) Control of plant development and gene expression by sugar signalling. Current Opinion in Plant Biology 8:93-102

Gonzalez-Lamothe R, Boyle P, Dulude A, Roy V, Lezin-Doumbou C et al. (2008) The transcriptional activator Pti4 is required for the recruitment of a repressosome nucleated by repressor SEBF at the potato PR-10a gene. The Plant Cell 20: 3136-3147

Gupta AK, Kaur N (2005) Sugar signalling and gene expression in relation to carbohydrate metabolism under abiotic stresses in plants. Journal of Biosciences 30:761-776

Harding K (2004) Genetic integrity of cryopreserved plant cells: A review. CryoLetters, 25: 3-22

Harding K, Johnston JW, Benson E (2009) Exploring the physiological basis of cryopreservation success and failure in clonally propagated in vitro crop plant germplasm. Agricultural and Food Science 18: 103-116

Heine-Dobbernack E, Kiesecker H, Schumacher HM (2008) Cryopreservation of dedifferentiated cell cultures. In: Reed BM (ed.) Plant Cryopreservation; A practical guide. Springer, New York

Hothorn T, Bretz F, Westfall P (2008) Simultaneous inference in general parametric models. Biometrical Journal 50: 346-363

Jain S, Srivastava S, Sarin NB, Kav NNV (2006) Proteomics reveals elevated levels of PR-10 proteins in saline-tolerant peanut (*Arachis hypogaea*) calli. Plant Physiology and Biochemistry 44: 253-259

Jelloulli N, Jouira HB, Skouri H, Ghorbel A, Gourgouri A, Mliki A (2008) Proteomic analysis of Tunisian grapevine cultivar Razegui under salt stress. Journal of Plant Physiology 165: 471-481

Kaczmarcyk A, Rokka VM, Keller ERJ (2011) Potato shoot tip cryopreservation. A review. Potato Research 54: 45-79

Kav NN., Srivastava S, Goonewardene L, Blade SF (2004) Proteome-level changes in the roots of *Pisum sativum* L. in response to salinity. Annals of Applied Biology 145: 217-230

Kaviani B (2011) Conservation of plant genetic resources by cryopreservation. Australian Journal of Crop Science 5: 778-800

Kishor PBK, Sangam S, Amrutha RN, Laxmi PS, Naidu KR et al. (2005) Regulation of proline biosynthesis, degradation, uptake and transport in higher plants: Its implications in plant growth and abiotic stress tolerance. Current Science 88: 424-438

Liu JJ, Ekramoddoullah AKM (2006) The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses. Physiological and Molecular Plant Pathology 68: 3-13

Lu ZW, Popova EV, Wu CH, Lee EJ, Hahn EJ et al. (2009) Cryopreservation of *Gingko biloba* cell culture: Effect of pretreatment with sucrose and ABA. CryoLetters 30: 232-243

Marcovics-Housley Z, Degano M, Lamba D, Roepenack-Lahaye E, Clemens S et al. (2003) Crystal structure of a hypoallergenic isoform of the major birch pollen allergen bet v 1 and its likely biological function as a plant steroid carrier. Journal of Molecular Biology 325: 123-133

Matton DP, Brisson N (1989) Cloning, expression and sequence conservation of pathogenesis related gene transcripts of potato. Molecular Plant-Microbe Interactions 2: 325-331

Menges M, Murray JAH (2004) Cryopreservation of transformed and wild-type *Arabidopsis* and *tobacco* cell suspension cultures. The Plant Journal 37: 635-644

Nicot N, Hausman JF, Hoffmann L, Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. Journal of Experimental Botany 56: 2907-2914

Ochatt SJ, Marconi PL, Radice S, Arnozis PA, Caso OH (1999) In vitro recurrent selection of potato: production and characterization of salt tolerant cell lines and plants. Plant Cell, Tissue and Organ Culture 55: 1-8

Panis B, Lambardi M (2006). Status of cryopreservation technologies in plants. In: Ruane J and Sonnino A (ed.) The role of biotechnology in exploring and protecting agricultural genetic resources. Food and Agriculture Organisation of the United Nations, Rome

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Research 29(9):e45

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Reed BM (2008) Plant cryopreservation: a practical guide. Springer, New York

Sabbah S, Tal M (1990) Development of callus and suspension cultures of potato resistant to NaCI and mannitol and their response to stress. Plant Cell, Tissue and Organ Culture 21:119-128

Salekdeh GH, Siopongco J, Wade LJ, Ghareyazie B, Bennet J (2002) Proteomic analysis of rice leaves during drought stress and recovery. Proteomics 2: 1131-1145

Sakai A, Kobayashi S, Oiyama I (1990) Cryopreservation of nucellar cells of navel orange (*Citrus sinensis* Osb. var. *brasiliensis* Tanaka) by vitrification. Plant Cell Reports 9: 30-33

Schäfer-Menuhr A, Schumacher HM (1997) Cryopreservation of potato cultivars – design of a method for routine application in genebanks. Acta Horticulturae 447: 477-482

Srivastava S, Fristensky B, Kav NNK (2004) Constitutive expression of a PR-10 protein enhances the germination of *Brassica napus* under saline conditions. Plant Cell Physiology 45: 1320-1324

Schaarschmidt F, Vaas L (2009) Analysis of trials with complex treatment structure using multiple contrast tests. *HortScience* 44: 188-195

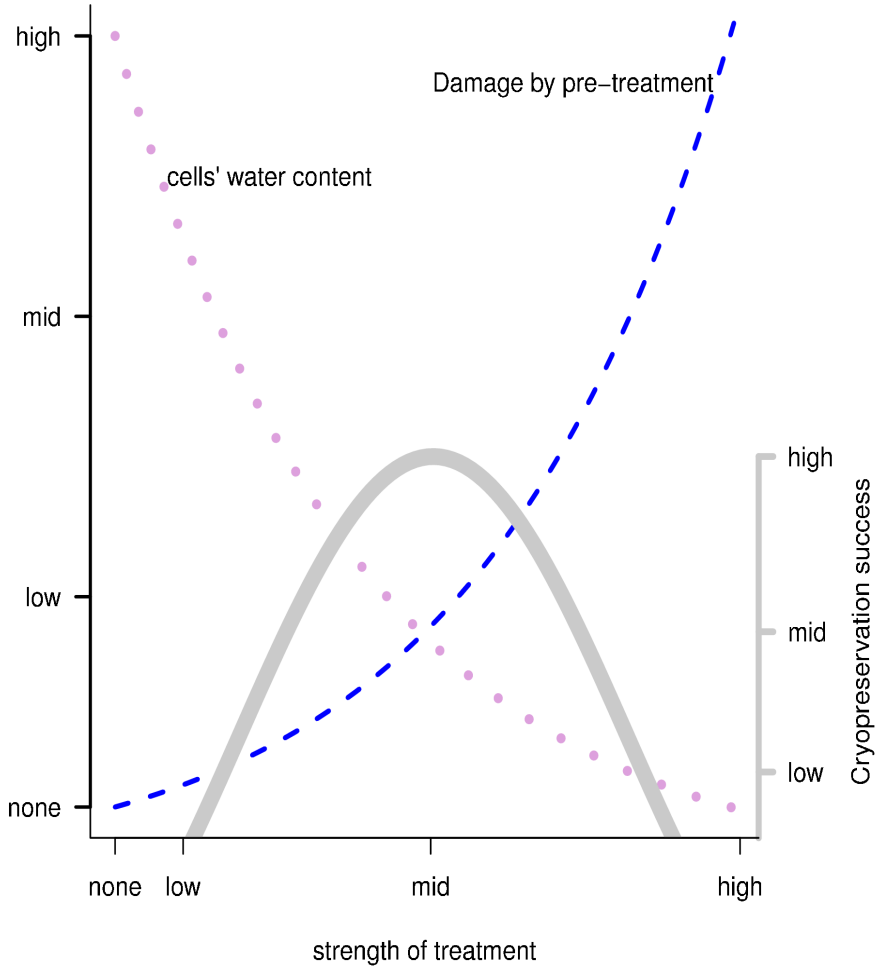
Swennen R, Carpentier S, Henry I, Vertommen A, Van den houwe I et al. (2011) From fundamental research discoveries to applications for banana improvement. Acta Horticulturae 897: 47-54

Volk GM (2010) Application of functional genomics and proteomics to plant cryopreservation. Current Genomics 11: 24-29

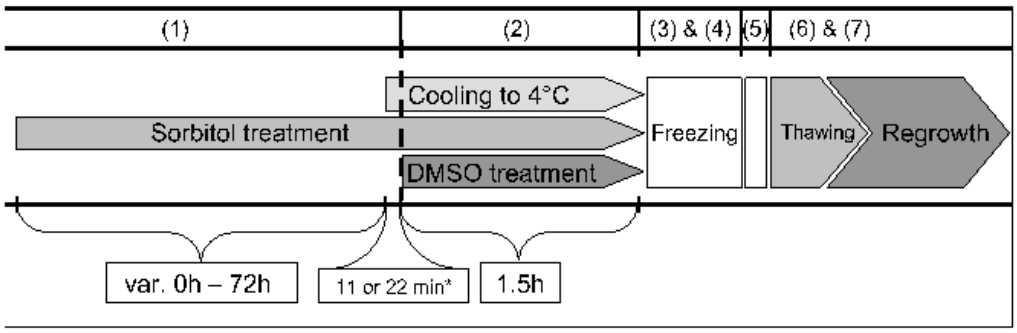
Withers LA and King PJ(1979) Proline: a novel cryoprotectant for the freeze preservation of cultured cells of *Zea mays* L. Plant Physiology 64: 675-678

Withers, LA and King, PJ (1980) A simple freezing-unit and routine cryopreservation method for plant cell cultures. CryoLetters 1: 213-220

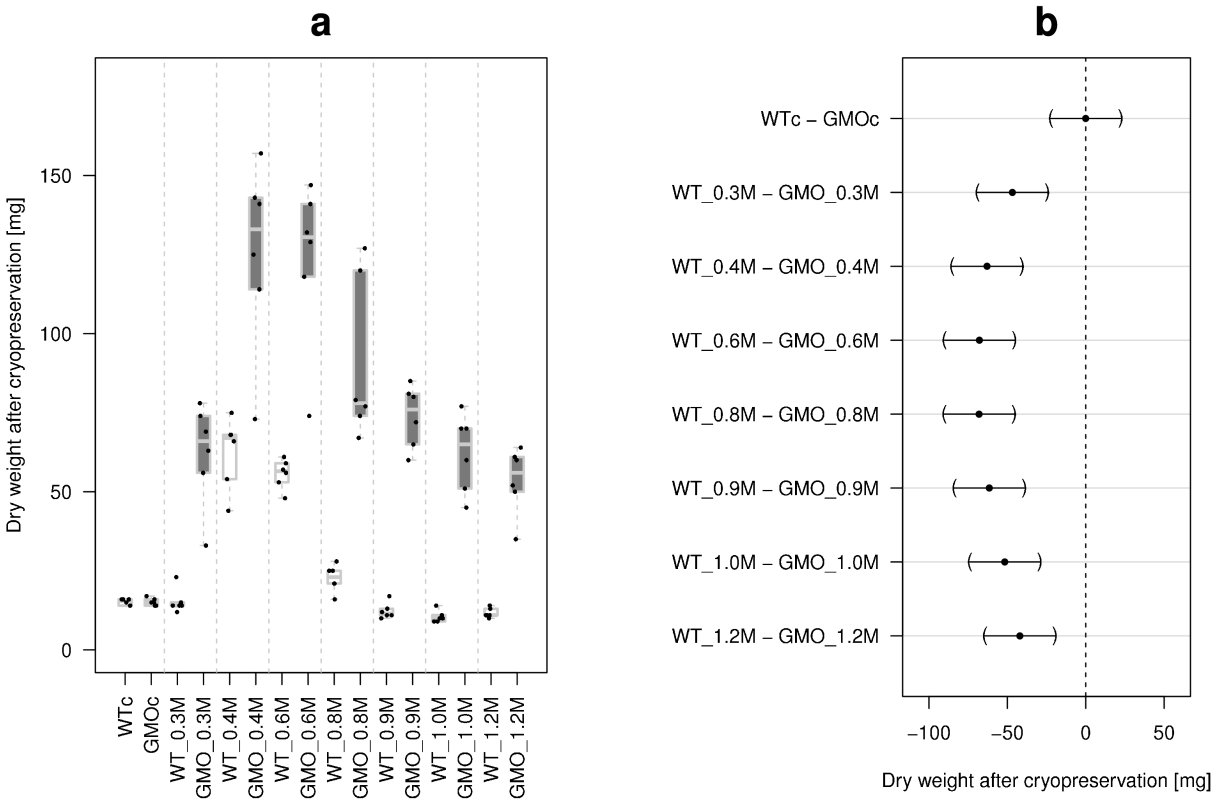
Zhou R, Cheng L, Dandekar AM (2006) Down-regulation of sorbitol dehydrogenase and up-regulation of sucrose synthase in shoot tips of the transgenic apple trees with decreased sorbitol synthesis. Journal of Experimental Botany 57: 3647-3657



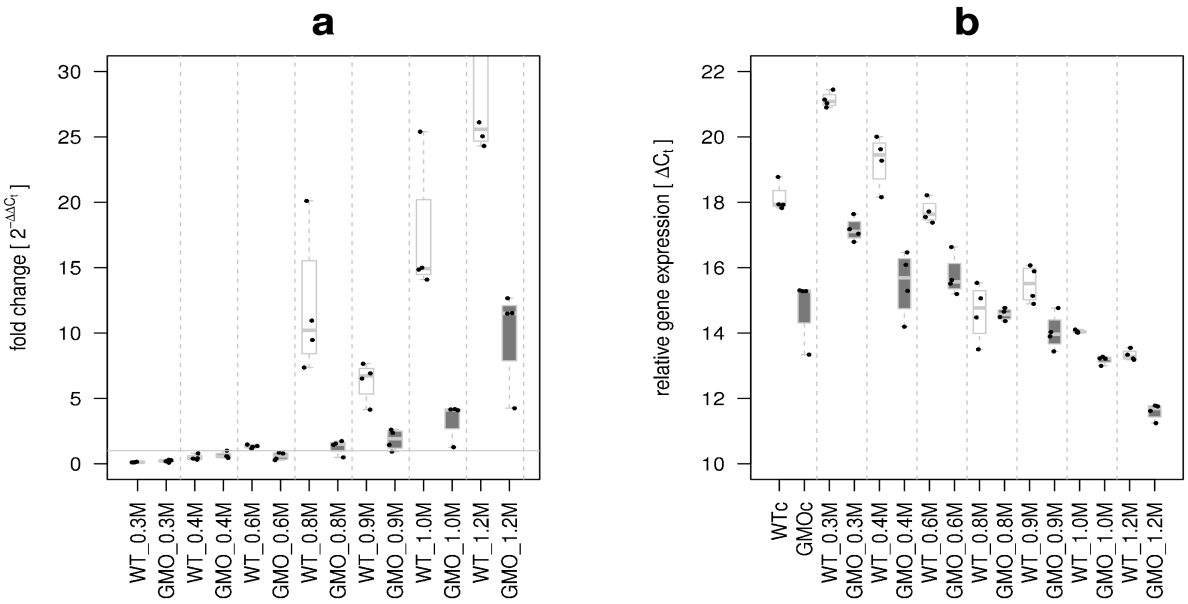
Chapter III, figure 1. Schematic relationship between cell dehydration (pink dotted line) and the cells' damages under increasing pretreatment strength (blue dashed line). The cryopreservation success (solid grey line) is highest where cells are sufficiently dehydrated to maximally circumvent ice crystal formation during freezing, but are not damaged too much.



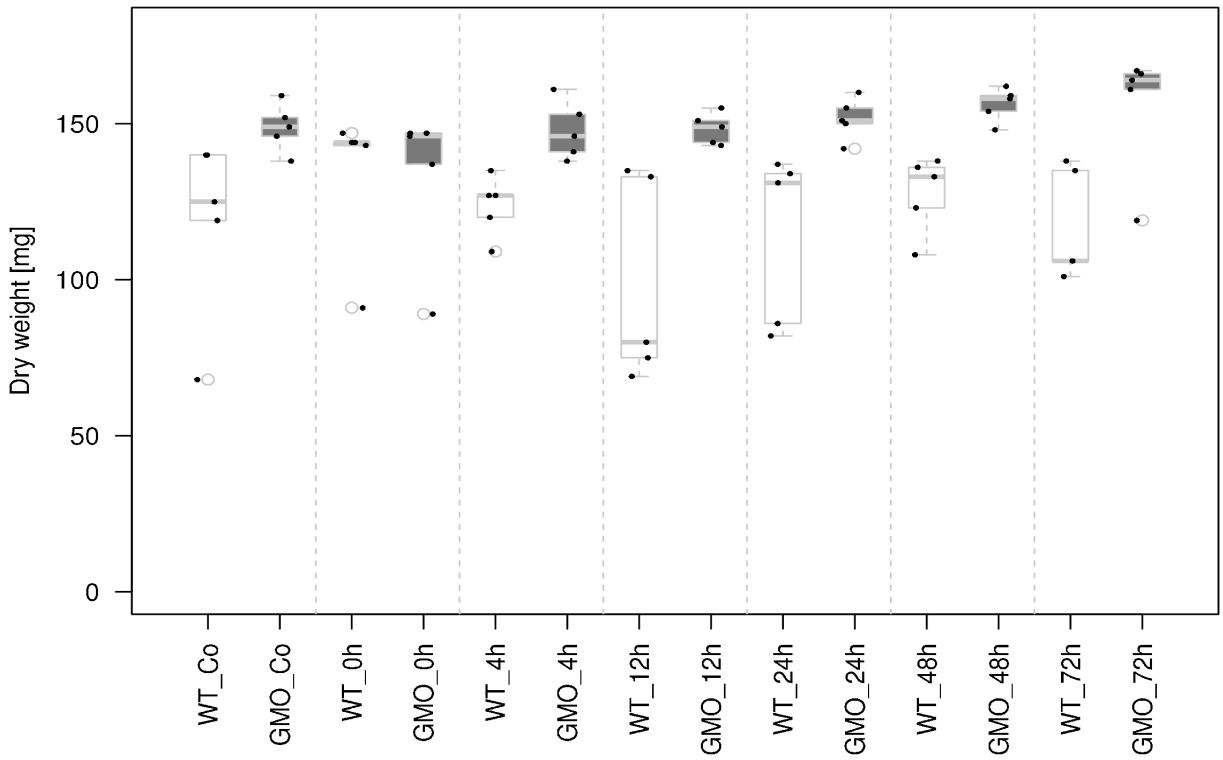
Chapter III, figure 2. Schedule of the cryopreservation procedure applied to potato suspension cultures. After a pretreatment step in medium supplemented with sorbitol, the cells were cooled to 4°C (*22min for the experiments using different sorbitol concentrations and 11min for the experiments using different incubation times of optimal sorbitol concentration) before the DMSO was added. After cells exposure to DMSO for 1.5h the freezing procedure, consisting of a slow cooling phase and the afterwards rapid immersion of samples in liquid nitrogen, was executed. After 24h storage in liquid nitrogen, thawing and regrowth were performed.



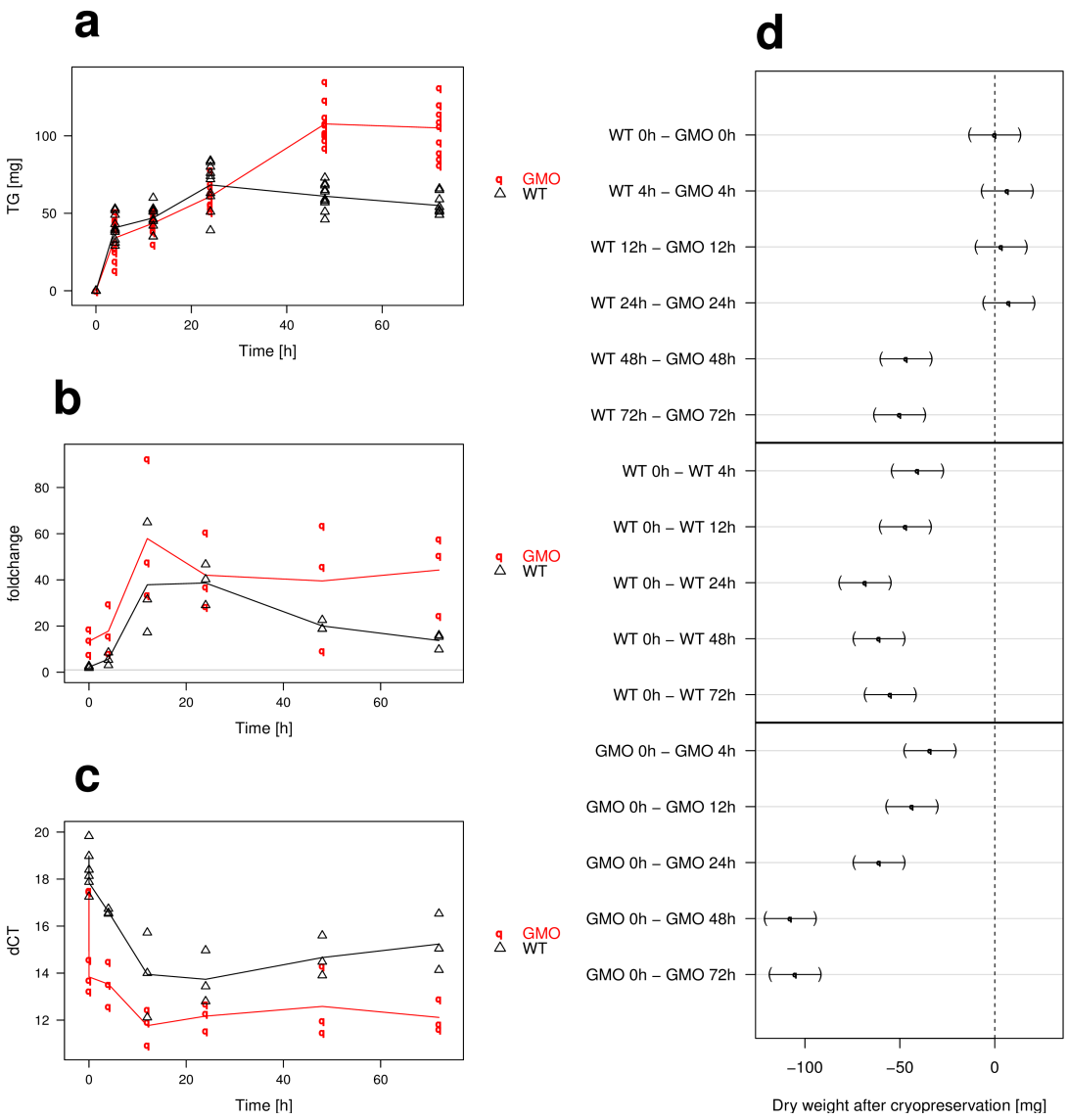
Chapter III, figure 3. Cryopreservation behaviour of wild type (WT) and transgenic (GMO) cell lines. (a) Cryopreservation success, measured as dry weight accumulation of cells re-grown after thawing determined under eight different sorbitol pretreatments. The results from six independent repetitions per condition (dots) are shown, complemented by a box plot. White boxes – WT, grey boxes – GMO. (b) mean values and their 95% confidence intervals for the simultaneous multiple comparisons of cryopreservation success of the genotypes for corresponding treatments.



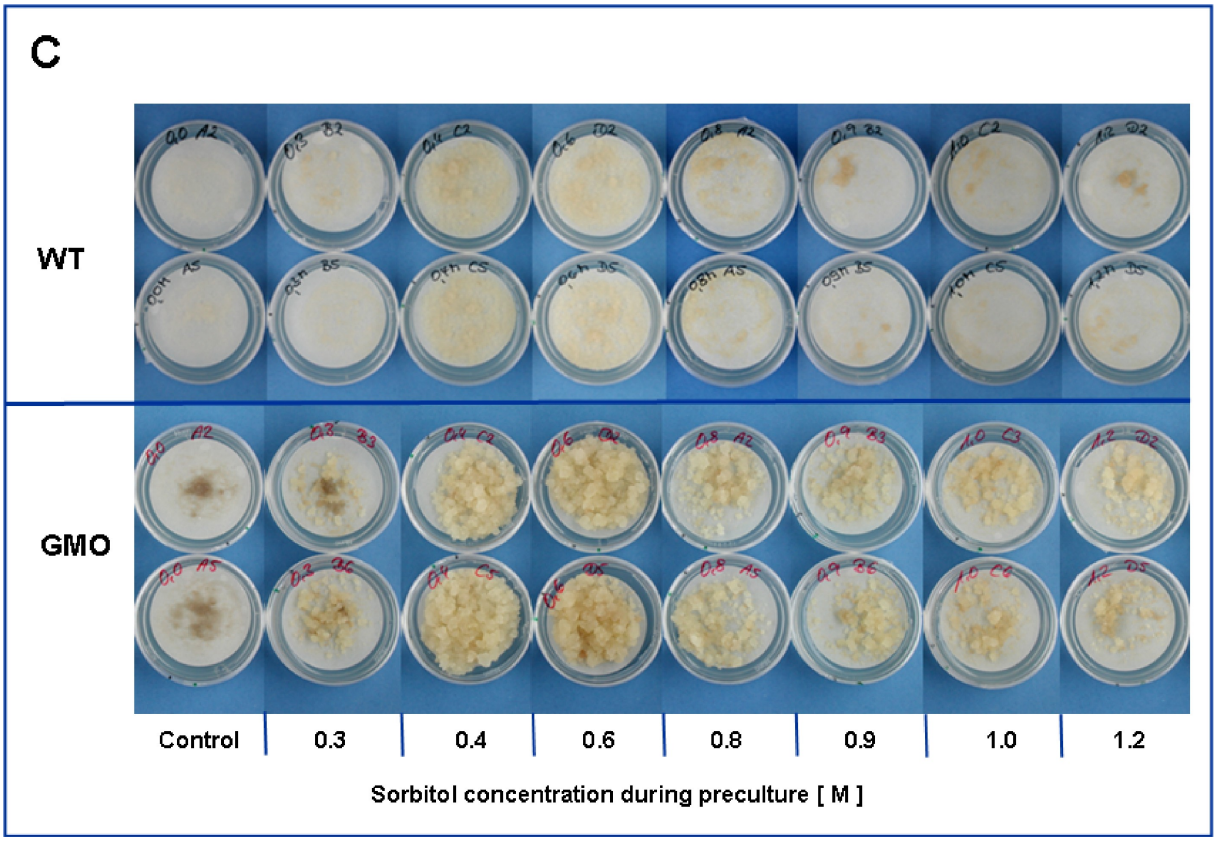
Chapter III, figure 4. (a) Results of *pr-10a* gene-expression fold change of cells treated with seven different sorbitol concentrations after the 48h pre-treatment interval. As control group cells were exposed to the normal 4X medium during the pregrowth time including the 1.5h long 5% DMSO supplementation. (b) Results of relative *pr-10a* expression analysis of cells treated with different sorbitol concentrations plus control group after the 48h pregrowth interval. Given are the expressions levels standardized to their *18S* rRNA threshold cycle (ΔCt). WT is given in white boxes and GVO in grey boxes. Note that smaller ΔCt values indicate a stronger expression (Pfaffl 2001).



Chapter III, figure 5. Diagram of regrowth behaviour measured as accumulated dry weight in mg of both the wild type (WT, white boxes) and the transgenic (GMO, grey boxes) cell line after cryopreservation pregrowth treatment without the freezing and thawing step. Given are the results of five independent repetitions per condition (dots), complemented by a box plot.



Chapter III, figure 6. Cryopreservation success and *pr-10a* expression at varying pretreatment durations but constant 0.6M sorbitol concentration. WT is coded with black lines and triangles, GMO with red lines and circles. (a) Dry-weight accumulation of cells after cryopreservation and re-grown for 15 days. The results from 16 independent repetitions per condition are shown. (b) Results on the *pr-10a* expression fold change over time. (c) Results of the analysis on the relative gene expression of *pr-10a* of cells treated with 0.6M sorbitol concentration after varying pretreatment durations. Note that smaller ΔCt values indicate a stronger expression (Pfaffl 2001). (d) Confidence intervals (95%) for simultaneous multiple comparisons of cryopreservation successes (data are from Fig. 6a). The means of the different genotypes for corresponding treatments are compared (upper panel), as well as the 0h-group against all others for WT (central panel) and GMO (lower panel). A 95% CI bar crossing the dotted vertical zero line indicates the absence of a significant difference.



Chapter III, figure 7. The effect of distinct sorbitol concentrations during a 48h long pretreatment phase on the quality of cryopreservability was investigated. Shown are cells after 15 days of regrowth with wild-type (WT) cells in the upper part and genetically modified (GMO) cells in the lower.

IV. Impacts of *pr-10a* overexpression at the molecular and the phenotypic level

Introduction

When plant cells are faced with osmotic or salt challenges, major changes in gene expression levels are intrinsic part of the drastic action triggering the physiological package of measures for stress response (Gaspar et al. 2002; Munns and Tester 2008). Because elevated *PR-10a* proteins are repeatedly found in salt and osmotically stressed entire plants as well as cell cultures (Kav et al. 2004; Jain et al. 2006; Jellouli et al. 2008; Jellouli et al. 2010) a role in stress perception or signal transduction has been postulated.

The first reports of the elicitor-induced appearance of mRNAs of the pathogenesis-related protein 10a (*pr-10a*, formerly known as STH-2) dates back more than twenty years ago (Marineau et al. 1987; Matton and Brisson 1989).

Meanwhile diverse studies reported *pr-10a* gene expression or protein abundance to be induced by several biotic and abiotic stressors in various plants, organs, tissues (Constable and Brisson 1995; Utriainen et al. 1998; Hashimoto et al. 2004; Mur et al 2004; Ukaji et al. 2004; Jelloulli et al. 2008) and developmental stages (Crowell et al. 1992; Walter et al. 1996). A detailed analysis of the expression pattern of the *pr-10a* gene in *Solanum tuberosum* plants of the cultivar Désirée revealed that no major organ exhibited constitutive expression (Constabel and Brisson 1995). Only the already known expression induction after infection, elicitor treatment, or, to a lower extent, after wounding could be confirmed with additional information about the magnitude of expression induction in vascular bundles, roots and leaves as well as in stigmas (Constabel and Brisson 1995).

Modulating the expression of *pr-10a* by genetic engineering yielded inconsistent results (reviewed in Liu and Ekramoddoullah 2006). Whereas some studies reported enhanced salt and/or osmotic tolerance due to *pr-10a* overexpression (Srivastava et al. 2004; El-Banna et al. 2010), the results in the context of pathogen attack easily interpretable. In potato plants overexpressing *pr-10a* no increase in resistance against either *Phytophthora infestans* or potato virus X (Constable et al. 1993) was observed, whereas in the legume *Medicago truncatula* silencing of PR-10-like proteins increased the tolerance against infection with *Aphanomyces euteiches* (Colditz et al. 2007).

Studies on the regulatory processes of *pr-10a* gene expression yielded more comprehensive results (Liu and Ekramoddoullah 2006), describing an interplay of a repressosome and an activator complex (Gonzales-Lamothe et al. 2008). Based on findings about the phosphorylation status of nuclear factor PBF-1 (Després et al. 1995) and the involvement of the single-stranded DNA binding factor Why1 (formerly PBF-2; Desveaux et al. 2000) in *pr-10a* gene activation and on observations of *pr-10a* repression by the single-stranded DNA binding protein SEBF (Boyle and Brisson 2001), it was hypothesized that the *pr-10a* gene has two activity states. In the inactivated state a repressosome, consisting of a heterodimeric SEBF-Pti4 complex, occupies the silencer element of the promoter (Gonzales-Lamothe et al. 2008). In order to become active, the repressosome has to be dismissed, thus allowing the recruitment of Why1 to the upstream elicitor response element in the promoter (Desveaux et al. 2000; Gonzales-Lamothe et al. 2008).

However, the mode of action of the PR-10a protein itself, as well as the pathways it could interfere with, are still an object of active research (Liu and Ekramoddoullah 2006; Gonzales-Lamothe et al. 2008). Beside reports about RNA hydrolysis (Bantignies et al. 2000; Zubini et al. 2009) the exploration of binding capacities of proteins of the *PR-10* family from different plants revealed high cytokinin affinity (Fujimoto et al. 1998; Gonneau et al. 2001), and other possible ligands like fatty acids, flavonoids (Fujimoto et al. 1998; Mogensen et al. 2002) or brassinosteroids (Markovic-Housely et al. 2003) were postulated. Additionally, possible crosstalks with hormone-signalling pathways (Genoud and Métraux 1999) as well as interaction with the mitogen-activated protein kinase cascades were reported (Xiong and Yang 2003). Further studies reported cryoprotective activity of *PR-10/Bet v 1* protein homologues in mulberry (Ukaji et al. 2004). To the best of our knowledge, however, none of the described PR-10a features were observed *in vivo*, thus the actual role of PR-10a proteins in living cells remains unclear (Van Loon and Van Strien 1999; Liu and Ekramoddoullah 2006).

Even though it is widely accepted that the gene does not encode a substantial new feature such as an ion-pump (Liu and Ekramoddoullah 2006), its overexpression can lead to enhanced tolerance in osmotic and/or salt challenges (see above). Because elevated PR-10a proteins are repeatedly found in salt and osmotically stressed entire plants (Jellouli et al. 2008), as well as cell cultures (Kav et al. 2004; Jain et al. 2006; Jellouli et al. 2010) a role in stress perception or signal transduction has been postulated, but also a LEA-like function protecting other proteins or cellular structures during drought or salinity was proposed (Liu and Ekramoddoullah 2006 and references therein).

Actually, none of the suggested functions is likely to be easily observable in tissues or even single cells. But the effect of *PR-10* protein abundances are ultimately displayed at the cellular level as cellular traits or phenotypes as apparent in growth curves (El-Banna et al. 2010). However, since plant cells' very first responses to salt or osmotic stress are rapid (within minutes to hours; Patterson 1979), the commonly applied but comparably slowly progressing growth measurements are not appropriate to determine the phenotype in these early phases.

As alternative to growth challenges, a more rapid and sensitive phenotypic parameter could be the determination of viability using assays based on triphenyltetrazolium chloride (TTC). Mitochondrial respiration leading to the production of NADH engenders a redox potential and flow of electrons to reduce a tetrazolium dye (Bochner and Savageau 1977), thereby producing insoluble purple coloured formazan . The more rapid the cellular respiration, i.e., the electron flow, the faster purple colour is formed (Bochner et al. 2001; Bochner 2009). TTC-based viability assays in plant cells utilize the difference between the colourless solution of TTC in water and the insoluble red formazan accumulating inside intact, living cells with a functional mitochondrial electron transport chain (Nachlas et al. 1960). Traditional applications of this assay mainly aimed at only qualitatively determining the viability of cells (dead versus alive). To this end, the accumulated formazan is extracted from the cells and its quantity is determined by measuring the absorption. The underlying assumption is that living cells produce the red formazan, which accumulates in the cells, while dead cells lose their membrane integrity, causing the formazan to escape into the surrounding medium (Duncan and Widholm 1990). Thus, the amount of absorption is proportional to the quantity of living cells.

Various studies suggested that quantitative evaluations are also feasible (Duncan and Widholm 1990; Zapata et al. 1991; Kurzbaum et a. 2010) and that additional information could be generated from quantifying both the finally accumulated amount of formazan as well as the amount during the course of its formation, i.e., the course of the respiration over time (Ruf and Brunner 2003; Duncan and Widholm 2004).

This principle of studying an organisms' behaviour on a longitudinal level was recently augmented by the introduction of a highly automated system measuring microbial respiration over time in a highly parallelized manner, the cells being faced with a large number of distinct physiological challenges (Bochner 2009). Since the successful application of this approach was reported for eukaryotes such as filamentous fungi (Hamer et al. 2001) and mammals

(Bochner et al. 2011), it can be postulated that this technique would permit the determination of phenotypes in a plant-cell system, too.

In this study, a potato wild-type cell line (*Solanum tuberosum* cv. Désirée) will be compared to two cell cultures homologously overexpressing the *pr-10a* gene. Firstly, a detailed analysis regarding the relative gene expression patterns of *pr-10a* as well as *sebf* and *pti4*, whose gene products are involved in the repressosome, provides insights at the molecular level into the response of heterotrophic cells to a variety of osmotic and salt-stress conditions. Secondly, to describe effects on a long term scale, we identify osmotic and salt-stress response also by a modification of the growth behaviour of the cells. Finally, to gather phenotypic information also about the very first phase of salt and osmotic stress response, we introduce an approach facilitating the measurement of respiration kinetics in heterotrophic plant cells. Its technical realisation will be discussed and the characteristics of the resulting type of data as well as requirements for their appropriate analysis will be figured out.

Materials and Methods

Plant material and osmotic challenge

As shown by Ali et al. (2010) dicistronic constructs achieve the successful coordinated co-expression of a physically independent target protein, here the *pr-10a*, providing a physiological trait, along with a reporter protein, a firefly luciferase. A non-embryogenic suspension culture of *Solanum tuberosum* cv. Désirée (DSMZ No. PC-1182) and two independently derived transgenic cell lines constitutively homologously overexpressing *pr-10a* (07-08-1 and 07-08-2, see El-Banna et al. 2010), were sub-cultured weekly by transferring 40 ml of suspension to 60 ml of fresh 4× medium (Gamborg et al. 1968) containing $2 mgl^{-1}$ 2,4-dichlorophenoxyacetic acid (2,4 D), $0.5 mgl^{-1}$ indole-3-acetic acid (IAA), $0.5 mgl^{-1}$ 1-naphtylacetic acid (NAA) and $0.4 mgl^{-1}$ kinetin, pH 5.6 and incubated in a 300ml Erlenmeyer flask on a gyratory shaker (TR- 250, Infors AG, Basel, Switzerland) with 50mm orbit (100 rpm) at 25°C. Figure 1 in the appendix to this chapter confirms the differential patterns of stable T-DNA integration in the transgenic cell lines by Southern blot analysis. In Figure 2 in the appendix of this chapter the time course of the corresponding relative expression levels of *pr-10a* (normalized against *18S* (Pfaffl et al. 2001) are given. Considering that smaller ΔCt values indicate higher expression levels, both transgenic cell lines showed substantially higher *pr-10a* expression levels from the beginning on.

To yield preculture cells for the separate experimental approaches (see below), cells were harvested from the logarithmic growth phase 3 days after subculturing. For that, cell mass was filtered off the medium through a Nylon net (100µm pore size, NeoLab, Heidelberg, Germany) using a Buchner funnel. The individual procedures for the different assays are described below.

Treatments consisting of 0.5M sorbitol and 0.32M NaCl supplements result in equiosmolal (800mOs/kg) concentrations of the stress agents, whereas the 0.16M NaCl yields only 500mOs/kg. Compared with the control 4× medium (200mOs/kg) the mild salt treatments (0.16M NaCl increases the osmotic pressure by a factor of 2.5, whereas the high salt and the sorbitol treatments lead to an increase by a factor of 4.

RNA isolation and quantitative real-time PCR

An inoculum of 15g preculture cell material (see above) was transferred into a sterile 300ml Erlenmeyer flask and filled up to 100ml with 4× medium supplemented with either 0.16M NaCl or 0.5M sorbitol or 0.32M NaCl.

Cell material was sampled after 3h, 4h, 9h, 10h, 24h, 48h and 72h incubation time. Medium was removed by filtering off the cells through a Nylon net (100µm pore size, NeoLab, Heidelberg, Germany) in a Buchner funnel. For each time point and treatment five independent flasks were inoculated and harvested independently, resulting in 420 individual samples. After harvesting the cell mass was immediately frozen in liquid nitrogen until further analysis.

The frozen cell material was ground under liquid nitrogen using a mortar and pistil. Total RNA was extracted from 100 mg fresh weight ground material with TriFast Gold (Peqlab, Erlangen, Germany) according to the manufacturer's instructions. cDNA synthesis was carried out with the RevertAid First Strand Synthesis kit (Fermentas, St. Leon Roth, Germany). For relative quantification of the *pr-10a* and luciferase mRNA by PCR, TaqMan probes against the *pr-10a* sequence (primers pr-10a129(f) 5'-TACACATGAAGCCACAAGCA-3', pr-10a129(r) 5'-ATGCTTCCATCTCCCTCAGT-3', probe: pr-10a129 5'-TCAAAGCTTTGGTTGTTGATGCTGA-3') and the luc sequence (primers luc141(f) 5'-TATGAACATTTCGCAGCCTA-3', luc141(r) 5'-ATCGACTGAAATCCCTGGTA-3', probe: luc141 5'-GTTTCCAAAAAGGGGTTGCAAA-3') were used.

For the relative quantification of the *sebf* and *pti4* mRNA by PCR TaqMan probes (primers ; (f) 5'-CCTTCTCCAATGGCTTCTTC-3', SEBF(r) 5'-GTTGTTTGGGAAGTGGGTTT-3', probe: SEBF(famtam) 5'-TCCCTCCATTTCCTTTCACTTACACCA-3' and Pti4(f) 5'-GGTTCAATGAAACGGAGAAGA-3', Pti4(r) 5'-GGACACCTGTCAATTGTTCG-3', Probe: Pti4(famtam) 5'-CCGTCACATTTCCGAACGGC-3') were used.

As internal standard, primers against the *18S* rRNA gene 18S138 (f) 5'-TAAAGGAATTGACGGAAGGG-3', 18S138 (r) 5'-CACCACCACCCATAGAATCA-3', probe: 18S138 5'-CGCAGGCTCCACTCCTGGTG-3' were used according to Nicot et al. (2005). Quantitative real-time PCR was performed on an Eppendorf Mastercycler ep realplex4 platform using the following program: 30s 95°C; [40× (95 °C, 5s; 60°C, 20s)].

To determine the effect of the osmotic challenges for each genotype separately, cells treated with the 4× medium served as the corresponding control group.

Dry-weight determination of cell material

For testing stress tolerance by growth, an inoculum of 1g of preculture cells (see above) per 100ml Erlenmeyer flasks was filled up with 50ml of 4× medium supplemented with either 0.16M NaCl or 0.5M sorbitol or 0.32M NaCl. The flasks were sealed with aluminium foil and incubated on a gyratory shaker (TR- 250, Infors AG, Basel, Switzerland) with 50mm orbit (100 rpm) at 25°C for up to 18 days.

Cell material was sampled every two days where five flasks were harvested independently for each treatment and dry weight was determined. To this end, the fresh cell material was transferred quantitatively into pre-weighed Petri dishes and dried to constant weight at 60°C for 72h. Petri dishes including dry cell matter were weighted again after drying. This experiment was repeated three times. For matters of convenience, below the data set of one of these repetitions is presented.

Respiration curves

For the respiration measurements, the 100ml Erlenmeyer flasks were set up analogously to the growth experiment with each 50ml medium additionally supplemented with 500µl of Biolog Redox Dye A® (100×) (Biolog Hayward CA). The flasks were sealed with aluminium foil and incubated on a gyratory shaker (TR-250, Infors AG, Basel, Switzerland) with 50mm orbit (100 rpm) at 25°C for up to eight hours in the dark. The formazan development was recorded after 2h, 4h, 6h, 8h and 24 h. For each treatment-genotype combination three flasks were

harvested independently. The flasks' content was transferred to 50ml PP tubes (Cellstar® Greiner bio-one) and cell material was collected by centrifugation with 4200g for 10 min. Supernatant was discarded. To stop the respiration activity of the cells, the pellet was immediately frozen at -80°C (Duncan and Widholm 1990). For extraction of the tetrazolium dye, the pellets were resuspended in 4ml ethanol (technical grade) and incubated in the dark over night at room temperature according (Ruf and Brunner 2003). By centrifugation (15min at 18000g) cell debris was removed and the absorption of the supernatant was determined photometrically at 520nm (TECAN Infinite M200 with 96er flat bottom PS Microplates (Greiner bio-one)). This experiment was repeated three times. For matters of convenience, below the dataset of one repetition is presented

As recommended by Duncan and Widholm (1990), the validity of the testing system was ascertained as following. Possible interactions between the used dye and the applied media were investigated; the correlation between formazan development in actively growing cells and samples killed by freezing was also determined. Furthermore, tests on background absorption due to extractable cellular pigments were performed, and the absorption spectra of the ethanol extracts were validated (figure 6 in the appendix to this chapter). Neither an interaction between the dye and the applied medium itself, nor false-positive formazan development in the here used *Solanum tuberosum* cv. Désirée cell cultures was observed (see Figure 7 in the appendix to this chapter).

Assessment of cell survival via respiration measurements

An inoculum of 1g preculture cells (see above) per flask was filled up with 50ml of 4× medium supplemented with 0.32M NaCl. The flasks were sealed with aluminium foil and incubated on a gyratory shaker (TR-250, Infors AG, Basel, Switzerland) with 50mm orbit (100 rpm) at 25°C. After 6d, 8d, 10d and 12d to three flasks per genotype 500µl of Biolog Redox Dye A® (100×) (Biolog Hayward CA) were added and incubated for additional 24h. The amount of formed formazan was determined as described above.

Statistical analysis

Calculation of both dCT and ddCT values was done using spreadsheets and the statistical software R (R Development Core Team 2010). Values were calculated as fold change $2^{-\Delta\Delta ct}$ for *pr-10a* standardized to the *18S* threshold cycle (ΔCt) with the differences between the treated cells and cells from control group. A fold change of 1 indicates no change in *pr-10a* expression caused by treatment (Pfaffl et al. 2001).

For graphical representation of experimental results basic plotting functions from the statistical software R (R Development Core Team 2010) together with the add-on package *lattice* (Sarkar 2008) were used.

For evaluation of growth curve results, a cell means model was set up describing the dry weight in dependency of the treatment (dry weight ~ treatment), whereas the treatment levels specify the combination of the genotype, the applied medium and the time of measurement, resulting in 108 levels.

Analogously, for evaluation of respiration curve results, a cell means model was set up, describing the absorption in dependency of the treatments (absorption ~ treatment), whereas the resulting 72 treatment levels consist specify the combination of the genotype, the applied medium and the time of measurement.

For comparisons of experimental group means simultaneous multiple-comparison procedures according to Hothorn et al. (2008) and Schaarschmidt and Vaas (2009) were performed with R (R Development Core Team 2010) using the above-mentioned linear cell-means models.

Results

Magnitude of *pr-10a* expression induction is sensitive to both the chemical compound and the strength of osmotic pressure

At the high osmotic pressure treatments of 0.5M sorbitol (Figure 1 A) or 0.32M NaCl (Figure 1 C), irrespective of the genotype, the pattern of gene-expression induction over time exhibited its maximum at 10h after start of treatment followed by a decline back to the low levels of the begin. In the wild-type cells, the magnitude of this gene-expression induction differed dramatically between the two osmotic treatments. Whereas the 0.5M sorbitol treatment induced *pr-10a* gene expression to a relative fold change of about 35 (Figure 1 A), while the equiosmolal 0.32mM NaCl treatment caused a much larger fold change of about 300 (Figure 1 C). Contrastingly, both transgenic cell lines exhibited much less gene-expression induction. Genotype 07-08-1 reacts with a *pr-10a* expression fold change of around 20 to the sorbitol treatment (Figure 1 A) and genotype 07-08-2 with one around 20. Compared to the corresponding control treatments under 0.32M NaCl (Figure 1 C), the transgenic cell lines do not even reach one third of the fold change the wild type cells display under the same conditions.

Surprisingly, the 0.16M NaCl treatment entailed a fold change of only five to ten in all the genotypes, indicating an only weak *pr-10a* gene-expression induction (Figure 1 B). However, genotype 07-08-1 showed a slight expression induction at the last time point after 72h stress treatment.

Interestingly, all genotypes react to the mild 0.16M NaCl and the 0.5M sorbitol treatments with similar rates of expression induction. Due to the generally higher expression level of the transgenic cell lines, they perform on different levels, but parallely to the wild type cells throughout the observed period of time (see Figure 2 in the appendix to this chapter). However, the 0.32M NaCl treatment entails such a severe *pr-10a* expression induction in the wild-type cells, that they reach the also heavily induced level of the transgenic cell lines (depicted in Figure 2 in the appendix to this chapter).

Fast induction of *pr-10a* expression regulators *sebf* and *pti4*

In analogy to the *pr-10a* expression monitoring, transcriptional levels of *sebf* and *pti4*, whose gene products are engaged in the heterodimeric SEBF-Pti4 repressosome complex (Gonzales-Lamothe et al. 2008), were tracked in the same experimental setup and displayed in Figure 2. The according relative expression levels of *sebf* and *pti4* (normalized against *18S* (Pfaffl et al. 2001) are provided in the appendix to this chapter (Figures 3 and 4, respectively).

Interestingly, in contrast to the *pr-10a* expression, for the regulator coding genes, both transgenic lines showed a far stronger induction than the wild type. Certainly, both the regulator coding genes showed expression patterns corresponding to that exhibited by *pr-10a* shown above, but mainly differing in the magnitude of the fold changes.

The high osmotic pressure of the 0.5M sorbitol treatment revealed a slight, about five-fold gene-expression induction for both regulators in the transgenic cell lines, but only during the first ten hours. Genotype 07-08-1 seemed to be more strongly reacting than 07-08-2. In contrast, the gene expression of the wild type remained rather constant.

Under the mild 0.16M NaCl regime, a similar pattern revealed a slight expression induction during the very first phase of stress treatment followed by decline to the non-induced level. Further, the comparably weak induction of *pr-10a* expression genotype 07-08-1 exhibited at the 72h time point (around tenfold see Figure 1 B) corresponded to an induction of both *sebf* and *pti4* at that time point (Figure 2).

During the first ten hours the severe 0.32M NaCl treatment generated expression induction comparable to that from the mild 0.16M NaCl treatment. But with increased incubation time,

07-08-1 showed persistently enhanced *sebf* gene-expression induction, as well as nearly no decline in *pti4* expression fold change. Contrastingly, wild-type cells and genotype 07-08-2 exhibited nearly equally low expression fold-change values throughout the observation period.

Mannopine-synthase promoter is induced by strong salt treatments

The luciferase expression clearly increases due to the severe 0.32M NaCl treatment, with even higher fold changes (about 15fold) in genotype 07-08-1 compared to about 8fold gene induction in genotype 07-08-2. Interestingly 07-08-1 also reacts to both the mild 0.16M NaCl and the 0.5M sorbitol treatment with about seven to ten fold *luc* gene expression induction. Genotype 07-08-2 does not exhibit appreciable expression induction under these treatments.

Furthermore, the pattern of luciferase induction is remarkably similar to that observed for *sebf* and *pti4* (Figure 2)

Growth behaviour is influenced by both the chemical compound and the strength of the osmotic pressure

To enable comprehensive insights into the growth behaviour, the data are presented as both grouped according to treatments (Figure 3A) and grouped according to individual genotypes (Figure 3B).

Under control conditions, growth starts from the beginning on, with very similar growth behaviours of the three genotypes (Figure 3A). In contrast, at 0.5M sorbitol the start of the growth is delayed by approximately 10 days, but then shows an exponential increase until a mean dry weight of around 600mg was reached after 18 days (Figure 3A). Moreover, at 0.5M sorbitol, the growth of the genotypes differ. Whereas WT and genotype 07-08-1 behave somewhat similarly until day 10, the WT then follows an exponentially increasing course, but 7-08-1 increases rather linearly. At day 18 the mean dry weight of 07-08-1 was around 200mg lower than the one of the wild type (Figure 3A and Figure 5 in the appendix to this chapter). Contrastingly, the genotype 07-08-2 started its growth with a shorter delay of only four to six days. Apart from slight irregularities in the growth, 07-08-2 increased its dry weight nearly parallel to the wild type's curve. At day 18 the mean dry weight of 07-08-2 was around 200mg higher than the one of the wild type (Figure 3.A and Figure 5 in the appendix to this chapter). Interestingly, though the time point of growth start is retarded at 0.5M sorbitol, both the WT and genotype 07-08-2 yielded after 18 days large amounts of cell mass, suggesting a shorter generation time from day 10 on compared to control conditions.

Also at the milder salt stress of 0.16M NaCl the three genotypes showed severe growth retardation in comparison to the control condition. The wild type showed a delay of about twelve days before growth started, albeit only weakly until a mean dry weight of 200mg was reached after eighteen days (Figure 3 A). Contrastingly, both transgenic cell lines started growth after four to six days and showed better growth than the wild-type cells. At day eighteen 07-08-1 and 07-08-2 reached mean dry weights of around 250mg and 280mg, respectively.

The 0.32M NaCl treatment, though being equiosmolal to the 0.5M sorbitol, apparently caused the largest stress effect, none of genotypes showed any growth (Figure 3A).

For all genotypes, the growth is more severely retarded by both NaCl treatments, also the mild 0.16M NaCl, than by the 0.5M sorbitol treatment (Figure 3B). Interestingly, both transgenic cell lines showed less growth retardation due to all three types of osmotic treatments than the wild type (Figure 3B, middle and lower panel). Besides these common patterns, all three genotpyes nevertheless show different response types to the osmotic treatment. In genotype 07-08-01 growth at control condition, 0.5M sorbitol and 0.16M NaCl are somewhat similar in the sense that the begin of growth is quite early (~ day 3-5) and that growth is rather linear, irrespective of slight differences in the steepness of growth increase at different treatments. In contrast, though the growth behavious of 07-08-02 is similar to 07-08-01 in the first ten days across all four growth conditions, genotype 07-08-02 shows a growth increase from day ten on under control conditions and also under 0.5M sorbitol. The growth of the WT is characterized by a strong sensitivity to mild salt conditions (0.16M NaCl), and in contrast to the two transgenic genotypes, also by a substantial difference in growth characteristic between control condition and sorbitol. Under control conditions, the WT starts rapidly to grow and to reach its plateau already after 12 days. In contrast, with 0.5M sorbitol, the growth beginn is strongly retarded until day ten, but then progresses strongly to yield a cell mass at day 18 even larger than under control conditions.

Cell respiration reacts on both the chemical compound and the strength of osmotic pressure

Similar to the characterization of growth curves, for increased insight, also the respiration kinetics are presented as both grouped according to treatment (Figure 4A) and grouped according to genotypes (Figure 4B).

Under all investigated conditions, active respiration can be detected from the beginning of the experiment on, with the first measurement being performed already after 2 hours (Figure 4A). Also, under all investigated conditions, the respiration kinetics were astonishingly parallel for both the wild type and the transgenic cell lines. This is especially true for the 0.5M sorbitol treatment (Figure 4A). Solely, after 24h the transgenic cell lines accumulated detectably lower formazan intensities under control and the mild 0.16M NaCl treatment (Figure 4A).

The grouping according to genotypes revealed another interesting feature (Figure 4 B). Within each genotype, the respiration kinetics in the first 6-8 hours for the control condition and the two salt treatments are highly similar. In contrast, from the beginning on, the 0.5M sorbitol treatment caused a substantially larger increase in formazan accumulation in all three genotypes (Figure 4 B). Only at advanced time points, after 24 hours, all three genotypes differ in their respiration kinetics at the different osmotic treatments (Figure 4 B). After 24h the wild type showed similarly slight increased respiration due to the salt treatment, while 07-08-1 reacted strongly to the severe 0.32M NaCl but only very slightly to the mild 0.16M NaCl treatment. In contrast, 07-08-2 reacted sensitive to applied osmotic pressure and type of compound, exhibiting individual respiration curves for each investigated treatment (Figure 4B).

Wild-type cells showed sustained respiration capacity under severe 0.32M NaCl

Although the 0.32mM NaCl condition completely impeded growth in all genotypes (Figure 3), the determination of respiration capacity surprisingly revealed sustained respiration capacities in the wild type cells, while both transgenic cell lines exhibit a relatively fast decline (Figure 6). Thereby, the respiration performance of transgenic line 07-08-2 remains always inferior to both that from genotype 07-08-1 and the wild type cells.

Discussion

In this study, suspended wild-type potato cell cultures (*Solanum tuberosum* cv. Désirée) and two *pr-10a* overexpressing lines were investigated regarding their response to both osmotic and salt challenges. The expression pattern over time of *pr-10a* and of *sebf* and *pti4*, both engaged in the *pr-10a*-repressosome complex (Gonzalez-Lamothe 2008) was determined in detail. Along with information about growth performance and respiration behaviour the impact of *pr-10a* overexpression in both transgenic lines on both long and short term phenotypes are evaluated. When intact plants are facing salt and/or drought stress they are

able to decelerate or modify the impact of this condition by complex interactions between different organs and tissues (Munns and Tester 2008; Gaspar et al. 2002). These confusing spatial and temporal complexities on the complete organism's level define the stress responses at the cellular level and thus make a precise exploration of cause-and-effect chains rather difficult (Munns and Tester 2008).

Contrastingly, heterotrophic, dedifferentiated suspended cell cultures allow for higher convenience since no such superordinate mechanisms against salt and/drought stress are hampering experimental access. Cells are surrounded by the medium and they experience equal conditions at the same time. This facilitates the control of stress homogeneity and the characterisation of cell behaviour independent from plant morphology. For example, it is possible to distinguish between the effect of ionic and non-ionic solutes supplementing the media and to measure their adverse effects by comparatively simple methods determining dry weight accumulation (growth) or viability using staining assays (Lutts et al. 2004).

Obviously, heterotrophic dedifferentiated cell cultures have a limited repertoire of phenotypes compared to whole plants. Beside the established relatively slow developing growth curves a feature measurement for short term cell responses is of interest. Therefore, a method for monitoring the cell respiration behaviour using a modified cell-viability assay was adopted in this work. In addition to the common growth curves, this measurement was shown to provide information about the very first phase of a stress response, which are unattainable by the relative slow growth-measurement approach.

In order to be able to discriminate between ionic or non-ionic effects an non-ionic osmotic challenge (0.5M sorbitol) and an ionic equiosmolal sodium-chloride (0.32M NaCl) treatment were chosen both representing the four fold (800Os/kg) osmotic pressure of the control treatment (200Os/kg). The treatment set was further augmented by a milder 0.16M NaCl treatment (resulting in 500Os/kg) producing about 2.5fold osmotic pressure compared to control treatment.

Sorbitol, a sugar-alcohol, indeed changes the osmotic pressure, but is not likely to alter the ionic regime of the medium (Gopal and Iwama 2007). Although it was shown to leak into cells to some extent (Hohl and Schopfer 1991) it does not negatively affect metabolic activity, since it does not damage protein hydration. Indeed, Gopal and Iwama (2007) recommended the use of sorbitol for the in-vitro induction of water stress when working with intact *in-vitro* potato plants. But when working with dedifferentiated cells, the specific features of the

applied compounds affect the cells directly. Possible beneficial effects of sorbitol used as an additional carbon source, as described by Duncan and Widholm (2004) are discussed below.

NaCl as an ionic, membrane-permeable substance causes ionic stress in the cytoplasma and has severe impact on intact protein hydration when attaining the cytoplasma in excess (Gaspar et al. 2002; Munns and Tester 2008). Nevertheless, both sorbitol and NaCl induced the *pr-10a* expression in all investigated genotypes in such a similar pattern, that a common stimulus has to be postulated (Figure 1).

In the wild type cells any applied treatment induced well-defined and timed *pr-10a* expression patterns over time (Figure 1). The magnitude of expression induction varied according to type and strength of the applied challenge, but its incidence over time was not altered (Figure 1). Although the 0.5M sorbitol and the 0.32M NaCl treatments exerted the same strength of osmotic pressure on the cells, the wild type's *pr-10a* gene expression induction was around ten times higher in the salt treatment (Figure 1).

While the wild-type cells' growth was at least partially impaired by all applied treatments, the growth retardations showed no coincidence with the prevailing osmotic pressure (Figure 3). Mainly the difference between the 0.32M NaCl, which generally inhibits growth, and the 0.5M sorbitol treatment, which seems to become more beneficial the longer the cells are treated, calls attention. Contrastingly, the respiration under control conditions, the mild 0.16M and the severe 0.32M NaCl treatment did not differ during the first eight hours. Only the 0.5M sorbitol treatment led to significantly increased respiration.

Both the transgenic cell lines showed expression patterns principally similar to wild type cells, even when they displayed much lower gene-expression fold changes under the high osmotic pressure treatments (Figure 1). But considering the constitutive *pr-10a* overexpression in transgenic cell lines, they exhibited a constantly higher amount of *pr-10a* transcripts than the wild type throughout the observation time (Figure 2 in the appendix to this chapter).

Interestingly both transgenic lines showed generally better growth performance under either osmotic treatment as compared to the wild-type cells, although they also exhibited substantial differences between each other (Figure 3). This difference between the transgenic lines was mirrored also in respiration measurements. That is, like in the wild type cells, respiration of 07-08-1 did not differ between control conditions, the mild 0.16M and the severe 0.32M NaCl treatment during the first eight hours, whereas genotype 07-08-2 exhibited differences in

respiration between these treatments. Only the 0.5M sorbitol treatment led to significantly increased respiration in both transgenic lines (Figure 4).

Since growth is initially inhibited but further on enhanced together with primarily enhanced respiration by the sorbitol treatment, it can be hypothesized that this sugar alcohol stresses the cells osmotically from the beginning on, but with sustained treatment they can make use of sorbitol as a source of energy. Similar phenomena were observed by Duncan and Widholm (2004) with *Zea mays* stressed with sorbitol treatments. Metabolomic analyses indicated a fundamental adjustment of the carbon-cycling apparatus (Ulanov and Widholm 2010) suggesting that sorbitol is used simply as an additional carbon source, after cells have adapted to the new osmotic regime. That is, carbon balance as a major integrator of plants' response to stress (Nielsen and Orcutt 1996), and thus the maintained ATP production throughout stress, is seen as the crucial point to facilitate tolerance mechanisms (Jacoby et al. 2011; Atkin and Macharel 2009). Nevertheless, the primary osmotic stress phase has to be managed and at least transgenic line 07-08-2 seems to take advantage from the *pr-10a* overexpression in this early phase, as this line is able to adapt faster (Figure 3).

A second interesting phenomenon is that the osmotic pressure of the mild 0.16M NaCl treatment is 2.5 fold compared to the control medium but induces the wild type's *pr-10a* expression only marginally (Figure 1 B), whereas the transgenic cell lines showed considerably higher *pr-10a* induction under this condition (Figure 1 B).

Since the (q)PCR-method did not discriminate between recombinant and endogenous *pr-10a* transcripts the origin of the transcripts in the transgenic lines could not be determined. However, the relative gene-expression levels revealed a generally higher amount of *pr-10a* transcripts in both transgenic genotypes (see Figure 2 in the appendix to this chapter) indicating a well working constant overexpression of *pr-10a*. Considering that both *pr-10a* overexpressing cell lines responded only slightly to the applied stresses (Figure 1), the impact of the, apparently still acting, endogeneous *pr-10a* genes is supposed to be marginal compared to the total of the *pr-10a* transcripts. Nevertheless, the comparably high *pr-10a* induction in the transgenic cell lines under the mild salt stress suggests an interaction with the transgenic construct, namely a possible induction effect of the salt treatment onto the mannopine-synthase promoter (p-MAS) (Langridge et al. 1989; Fox et al. 1992) controlling the transgenic construct (Ali et al. 2010). The verification of the expression pattern of the luciferase, which is physically linked to the transgenic *pr-10a*, revealed supportingly a relatively strong, ten- to 15-fold induction due to the high 0.32M NaCl treatment and a comparably low, but not

negligible induction by sorbitol and the mild 0.16M NaCl treatment. Because the endogenous *pr-10a* responds with high gene-expression induction to the high osmotic pressure treatments, the enhanced expression of the transgene is not detectable, but under the mild 0.16M NaCl treatment this expression enhancement is not covered.

In the same experimental set-up, the expression patterns of *sebf* and *pti4* were monitored in order to elucidate whether the abundances of these regulator proteins are controlled by altered expression patterns under different stress regimes. Considering the expression induction of *sebf* and *pti4* in both the transgenic lines together with the expression-induction patterns of the *luc* transcripts, an interesting manifest coincidence could be obtained. That is, with the salinity-dependent induction of the p-MAS the regulators *sebf* and *pti4* exhibited a well-structured expression pattern. Since the wild-type cells altered neither the *sebf* nor *pti4* gene expression appreciably under the sorbitol or the equiosmolal 0.32M NaCl treatments, but both transgenic lines did, it is supposed that the expression induction in the transgenic lines is an effect from the transgenes.

Forming a heterodimeric SEBF-Pti4 complex, the two proteins were shown to disable the expression of *pr-10a* via the occupation of the silencer element of the promoter (Gonzales-Lamothe et al. 2008). For the activation of *pr-10a* gene expression, the complex has to dissociate from the DNA, and the proteins are hypothesized to also dissociate from each other (Gonzales-Lamothe et al. 2008).

Although various details about the transcriptional and posttranscriptional regulation of Pti4 (Gu et al. 2000; Mysore et al. 2002; Wu et al. 2003) and the determining features of the SEBF protein (Matton et al. 1993; Despres et al. 1995; Boyle and Brisson 2001) are known, the data at hand do not permit an unambiguous interpretation. However, considering the findings of Gu et al. (2002) indicating that, among others *pti*-genes, *pti4* activates the expression of a wide array of pathogenesis related genes and plays important and distinct roles in plant defence mechanisms, it has to be argued that due to unknown feedback as well as forward-loop regulations the *pr-10a* overexpression disturbed the adjustment of plant-defence gene regulation via modifying the *pti4* and *sebf* expression levels. Further research is necessary for investigation of these interaction mechanisms and their impact on the character of the cells' response to stresses.

This study supports the view that respiration measurements can provide valuable information especially during the very first phase of stress exposure, i.e. a short period of time in which the determination of growth reactions is usually infeasible (Figure 5). Furthermore, this

technique is likely to facilitate studies on the metabolic activities of cells under even severe stress conditions or other adverse circumstances (Lutts et al. 2004), as we have also shown here (Figure 6), respiration can occur independently of cell growth (Patterson 1979; Chapin 1991). For the data at hand it indicates that the *pr-10a* overexpression at least for the here investigated severe 0.32M NaCl stress is not exclusively beneficial and thus provides interesting information about the impact of artificially enhanced presence of *pr-10a* on cell physiology.

Unfortunately, the purple-coloured formazan accumulates inside the cells which makes the procedure of respiration measurement over time cost- and labour-intensive and thus does not yet facilitate a high-throughput approach in analogy to the Biolog PM technique (Bochner et al. 2011). Thus further research should aim at the discovery of tetrazolium formulations which can escape the cells and thus simplify the process of measurement by superseding the ethanol-extraction step.

The underlying assumption for the respiration measurements is that electron flow through the mitochondrial respiratory electron chain, which causes the formazan formation, automatically leads to ATP synthesis (Bochner 2009; Bochner et al. 2011). Since plant mitochondrial respirator chain can be by-passed using the non-energy conserving alternative oxidase pathway (reviewed in Millar et al. 2011), it has been hypothesized that the formed purple colour in TTC based viability assays must not automatically correspond to ATP production rates. However, experimental findings concerning the correlation of ATP production and formazan based respiration assays are ambiguous (Duncan and Widholm 2004, Kurzbaum et al. 2010, Galle et al. 2010), thus more research on this is necessary.

Moreover, data analysis should be reconsidered in advance, because with the improvement of measurement techniques the longitudinal character of the output data will become more important. Thus, data complexity will increase, but contain additional information coded in the shape characteristics of the curves (Brisbin et al. 1987). These curve features can unravel fundamental differences or similarities in the respiration behaviours under distinct treatments, which may be identifiable only by curve comparisons rather than the simpler comparisons of means.

Literatur

Ali Z, Schumacher HM, Heine-Dobbernack E, El-Banna A, Hafeez FY et al. (2010) Dicistronic binary vector system – A versatile tool for gene expression studies in cell cultures and plants. Journal of Biotechnology 145:9-16

Bantignies B, Seguin J, Muzac I, Dedaldechamp F, Gulick P et al. (2000) Direct evidence for ribonucleolytic activity of a PR-10-like protein from white lupin roots. Plant Molecular Biology 42: 871-881

Bochner BR and Savageau MA (1977) Generalized indicator plate for genetic, metabolic, and taxonomic studies with microorganisms. Applied and Environmental Microbiology 33: 434-444

Bochner BR, Gadzinski P, Panomitros E (2001) Phenotype MicroArrays for high throughput phenotypic testing and assay of gene function. Genome Research 11: 1246-1255

Bochner BR (2003) New technologies to assess genotype–phenotype relationships. Nature Reviews Genetics 4: 309-314

Bochner BR (2009) Global phenotypic characterization of bacteria. FEMS Microbiology Reviews 33: 191-205

Bochner BR, Siri M, Huang RH, Noble S, Lei XH et al. (2011) Assay of the multiple energy-producing pathways of mammalian cells. PLoS ONE 6: e18147

Boyle B, Brisson N (2001) Repression of the defense gene PR-10a by the single-stranded DNA binding protein SEBF. The Plant Cell 13: 2525-2537

Brisbin IL, Collins CT, White GC, McCallum DA (1987) A new paradigm for the analysis and interpretation of growth data: the shape of things to come. The Auk 104: 434-444

Chapin FS (1991) Integrated responses of plants to stress. A centralized system of physiological responses. BioScience 40: 29-31

Colditz F, Niehaus K, Kajinski F (2007) Silencing of PR-10-like proteins in *Medicago truncatula* results in an antagonistic induction of other PR proteins an in an increased tolerance upon infection with the oomycete *Aphanomyces euteiches*. Planta 226: 57-71

Constabel CP, Bertrand C, Brisson N (1993) Transgenic potato plants overexpressing the pathogenesis-related STH-2 gene show unaltered susceptibility to *Phytophthora infestans* and potato virus X. Plant Molecular Biology 22: 775-782

Constabel CP, Brisson N (1995) Stigma- and vascular-specific expression of the pr-10a gene of potato: a novel pattern of expression of a pathogenesis-related gene. Molecular Plant-Microbe Interactions 8: 104-1138

Crowell DN, John ME, Russel D, Amasona R (1992) Characterization of a stressinduced, developmentally regulated gene family from soybean. Plant Molecular Biology 18, 459-466

Després C, Subramaniam R, Matton DP, Brisson N (1995) The activation of the potato PR-10a gene requires the phosphorylation of the nuclear factor PBF-1. The Plant Cell 7: 589-598

Desveaux D, Després C, Joyeux A, Subramaniam R, Brisson N (2000) PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. The Plant Cell 12:1477-1489

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12, 13-15

Duncan DR, Widholm JM (1990) Measurement of viability suitable for plant tissue cultures. In: Polland JW, Walker JM (ed.) Methods in Molecular Biology Vol. 6 Plant Cell and Tissue Culture. The Humana Press, New Jersey

Duncan DR, Widholm JM (2004) Osmotic induced stimulation of the reduction of the viability dye 2,3,5-triphenyltetrazolium chloride by maize roots and callus cultures. Journal of Plant Physiology 161: 397-403

El-Banna A, Hajirezaei MR, Wissing J, Ali Z, Vaas L et al. (2010) Over-expression of PR-10a leads to increased salt and osmotic tolerance in potato cell cultures. Journal of Biotechnology 150: 277-287

Fox PC, Vasil V, Vasil IK, Gurley WB (1992) Multiple ocs-like elements required for efficient transcription of the mannopine synthase gene of T-DNA in maize protoplasts. Plant Molecular Biology 20: 219-233

Fujimoto Y, Nagata R, Fukasawa H, Yano K, Azuma M, Iida A, Sugimoto S, Shudo K, Hashimoto Y (1998) Purification and cDNA cloning of cytokininspecific binding protein from mung bean (*Vigna radiata*). European Journal of Biochemistry 258: 794-802

Galle A, Florez-Sarasal I, Thameur A, de Peape R, Flexas J, Ribas-Carbo M (2010) Effects of drought stress and subsequent rewatering on photosynthetic and respiratory pathways in *Nicotiana sylvestris* wild type and the mitochondrial complex I-deficient CMSII mutant. Journal of Experimental Botany 61: 765-775

Gamborg OL, Miller RA, Ojima K (1968) Nutrient requirements of suspension cultures of soybean root cells. Exp. Cell Res. 50: 151-158

Gaspar T, Franck T, Bisbis B, Kevers C, Jouve L, Hausman JF, Dommes J (2002) Concepts in plant stress physiology. Application to plant tissue cultures. Plant Growth Regulation 37: 263-285

Genoud T, Metraux JP (1999) Crosstalk in plant cell signaling: structure and function of the genetic network. Trends in Plant Science 4: 503-507

Gonneau M, Pagant S, Brun F, Laloue M (2001) Photoaffinity labelling with the cytokinin agonist acido-CPPU of a 34 kDa peptide of the intracellular pathogenesis-related protein family in the moss *Physcomitrella patens*. Plant Molecular Biology 46: 539-548

Gonzalez-Lamothe R, Boyle P, Dulude A, Roy V, Lezin-Doumbou C et al. (2008) The transcriptional activator Pti4 is required for the recruitment of a repressosome nucleated by repressor SEBF at the potato PR-10a gene. The Plant Cell 20: 3136-3147

Gopal J, Iwamain K (2007) in vitro screening of potato against water-stress mediated through sorbitol and polyethylene glycol. Plant Cell Reports 26: 693-700

Gupta AK, Kaur N (2005) Sugar signalling and gene expression in relation to carbohydrate metabolism under abiotic stresses in plants. Journal of Biosciences 30: 761-776

Gu YQ, Yang C, Thara VK, Zhou J, Martin GB (2000) Pti4 is induced by ethylene and salicylic acid, and its product is phosphorylated by the Pto kinase. Plant Cell 12: 771-786

Gu YQ, Wildermuth MC, Chakravarthy S, Loh YT, Yang C et al. (2002) Tomato transcription factors Pti4, Pti5 and Pti6 activate defence responses when expressed in *Arabidopsis*. Plant Cell 14: 817-831

Hamer L, Adachi K, Montenegro-Chamorro MV, Tanzer MM, Mahanty SK et al. (2001) Gene discovery and gene function assignment in filamentous fungi. Proceedings of the National Academy of Science USA 98: 5110-5115

Hashimoto M, Kisseleva L, Sawa S, Furukawa T, Komatsu S et al. (2004) A novel rice PR10 protein, RSOsPR10, specifically induced in roots by biotic and abiotic stresses, possibly via the Jasmonic Acid signaling pathway. Plant Cell Physiology 45: 550-559

Hohl M, Schopfer P (1991) Water relations of growing maize coleoptiles. Comparison between mannitol and polyethylene glycol 6000 as external osmotica for adjusting turgor pressure. Plant Physiology 95: 716–722

Jain S, Srivastava S, Sarin NB, Kav NNV (2006) Proteomics reveals elevated levels of PR10 proteins in saline-tolerant peanut (*Arachis hypogaea*) calli. Plant Physiology and Biochemistry 44: 253-259

Jellouli N, Ben Jouira H, Skouri H, Ghorbel A, Gourgouri A et al. (2008) Proteomic analysis of Tunisian grapevine cultivar Razegui under salt stress. Journal of Plant Physiology 165: 471-481

Jellouli N, Jouira BH, Daldoul S, Chenennaoui S, Ghorbel A et al. (2010) Proteomic and transcriptomic analysis of grapevine PR10 expression during salt stress and functional characterization in yeast. Plant Molecular Biology Reporter 28: 1-8

Kav NNV, Srivastava S, Goonewardene L, Blade SF (2004) Proteome-level changes in the roots of *Pisum sativum* in response to salinity. Annals of Applied Biology 145: 217-230

Kurzbaum E, Kirzhner F, Armon R (2010) A simple method for dehydrogenase activity visualization of intact plant roots grown in soilless cultures using tetrazolium violet. Plant Root 4: 12-16

Langridge WH, Fitzgerald KJ, Koncz C, Schell J, Szalay AA (1989) Dual promoter of *Agrobacterium tumefaciens* mannopine synthase genes is regulated by plant growth hormones. Proceedings of the National Academy of Science USA 86, 3219-3223

Liu JJ, Ekramoddoullah AKM (2006) The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses. Physiological and Molecular Plant Pathology 68: 3-13

Lutts S, Almansouri M, Kinet JM (2004) Salinity and water stress have contrasting effects on the relationship between growth and cell viability during and after stress exposure in durum wheat callus. Plant Science 167: 9-18

Nachlas MM, Margulies SI, Seligman AM (1960) Sites of electron transfer to tetrazolium salts in the succinoxidase system. Journal of Biological Chemistry 235: 2739-2743

Nicot N, Hausman JF, Hoffmann L, Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. Journal of Experimental Botany 56: 2907-2914

Nilsen E, Orcutt DM (1996) The physiology of plants under stress – Abiotic factors. John Wiley and Sons, Inc, New York

Marineau C, Matton DP, Brisson N (1987) Differential accumulation of potato tuber mRNA's during the hypersensitive response induced by arachidonic acid elicitor. Plant Molecular Biology 9: 335-342

Markovic-Housely Z, Degano M, Lamba D, von Roepenack-Lahaye E, Clemens S et al. (2003) Crystal structure of a hypoallergenic isoforms of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier. Journal of Molecular Biology 325, 123-133

Matton P, Brisson N (1989) Cloning, Expression and Sequence Conservation of Pathogenesis related Gene Transcripts of Potato. Molecular Plant-Microbe Interactions 2: 326-331

Mayr E (1997) The objects of selection. Proceedings of the Natinal Academy of Science USA 94: 2091-2094

Millar AH,Whelan J, Soole KL, Day DA (2011)Organization and Regulation of Mitochondrial Respirationin Plants. Annual Review of Plant Biology 62: 79-104

Mogensen JE, Wimmer R, Larsen JN, Spangfort MD (2002) The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands. Journal of Biological Chemistry 277, 684-692

Mur LAJ, Sturgess FJ, Farrell GG, Draper J (2004) The AoPR10 promoter and certain endogenous PR10 genes respond to oxidative signals in *Arabidopsis*. Molecular Plant Pathology 5: 435-451

Munns R, Tester M (2008) Mechanisms of Salinity Tolerance. Annual Review of Plant Biology 59: 651-681

Mysore KS, Crasta OR, Tuori RP, Folkers O, Swirsky PB et al. (2002) Comprehensive transcript profiling of Pto and Prf-mediated host defence responses to infection by *Pseudomonas syringae* pv. tomato. Plant Journal 32: 299-315

Papin JA, Reed JL, Palsson BO (2004) Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. TRENDS in Biochemical Sciences 29: 641-647

Patterson MK (1979) Measurements of growth and viability of cells in culture. Methods in Enzymology 58: 141-152

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Research 29: e45

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Ruf M, Brunner J (2003) Vitality of tree fine roots: Reevaluation of the tetrazolium test. Tree Physiology 23: 257-263

Sarkar D (2008) Lattice: Multivariate Data Visualization with R. Springer, New York

Schaarschmidt F, Vaas L (2009) Analysis of Trials with Complex Treatment Structure Using Multiple Contrast Tests. HortScience 44: 188-195

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology 98: 503-517

Srivastava S, Fristensky B, Kav NNV (2004) Constitutive expression of a PR10 protein enhances the germination of *Brassica napus* under saline conditions. Plant Cell Physiology 45: 1320-1324

Ukaji N, Kuwabara C, Takezawa D, Arakawa K, Fujikawa S (2004) Accumulation of pathogenesis-related (PR) 10/Bet v 1 protein homologues in mulberry (*Morus bombycis* Koidz.) during winter. Plant, Cell and Environment 27: 1112-1121

Ulanov A, Widholm JM (2010) Metabolic profiling to determine the cause of the increased triphenyltetrazolium chloride reduction in mannitol-treated maize callus. Journal of Plant Physiology 167: 1423-1431

Utriainen M, Kokko H, Auriola S, Sarrazin O, Kärenlampi S (1998) PR-10 protein is induced by copper stress in roots and leaves of a Cu/Zn tolerant clone of birch, *Betula pendula*. Plant, Cell and Environment 21: 821-828

Van Loon LC, Van Strien EA (1999) The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiological and Molecular Plant Pathology 55: 85-97

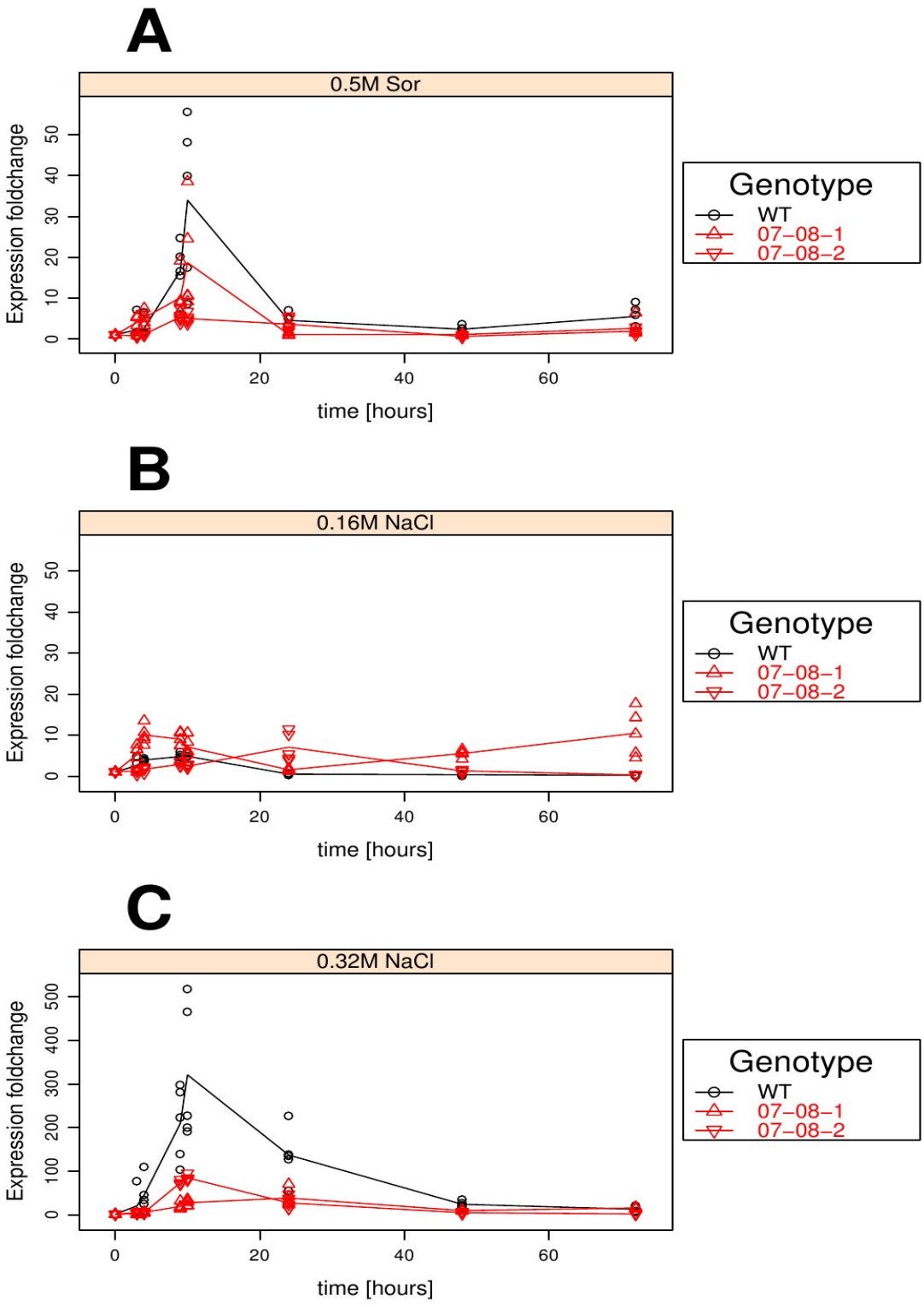
Walter MH, Liu JW, Wünn J, Hess D (1996) Bean ribonuclease-like pathogenesis-related protein gene (Ypr10) display complex patterns of developmental, dark-induced and exogenous-stimulus-dependent expression. European Journal of Biochemistry 239: 281-293

Wu K, Tian L, Hollingworth J, Brown DC, Miki B (2002) Functional analysis of tomato Pti4 in *Arabidopsis*. Plant Physiology 128: 30-37

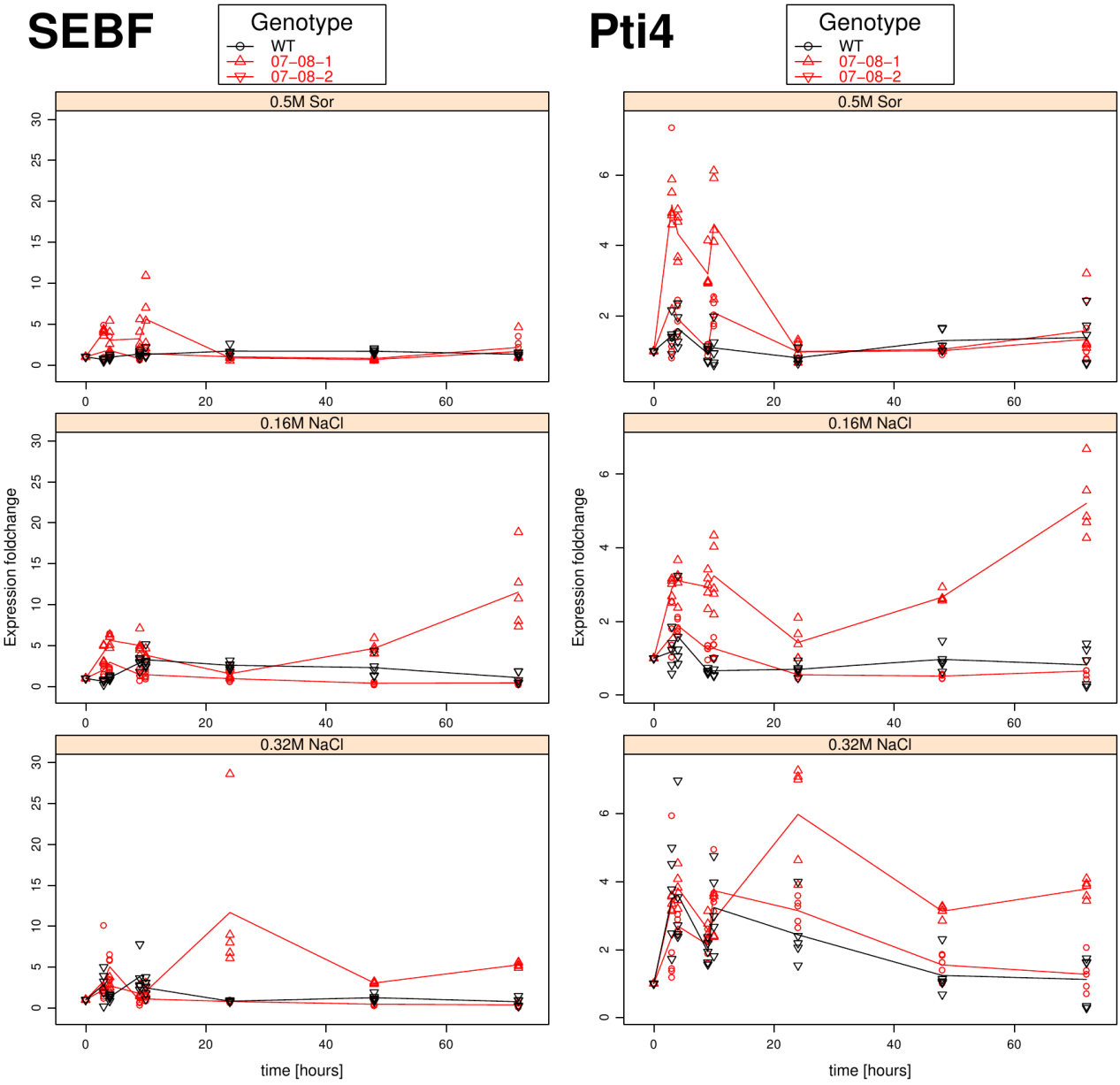
Xiong L, Yang Y (2003) Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acid–inducible mitogen-activated protein kinase. The Plant Cell 15: 745-759

Zapata JM, Salinas C, Calderon AA, Munoz R, Barcelo AR (1991) Reduction of 2,3,5-triphenyltetrazolium chloride by the KCN-insensitive, salicylhydroxamic acid-sensitive alternative respiratory pathway of mitochondria from cultured grapevine cells. Plant Cell Reports 10: 579-582

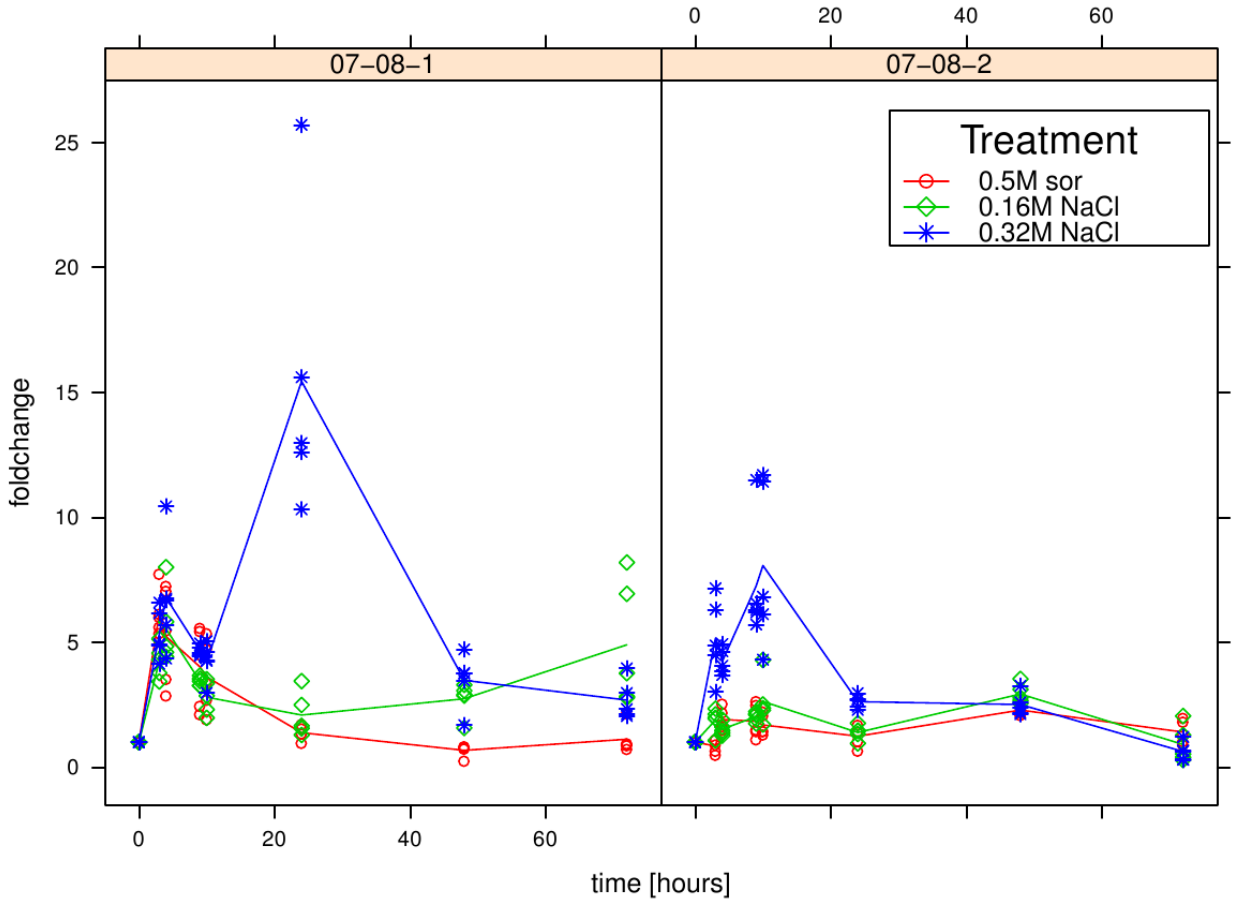
Zubini P, Zambelli B, Musiani F, Ciurli S, Bertolini P et al. (2009) The RNA hydrolysis and the cytokinin binding activities of PR-10 proteins are differently performed by two isoforms of the Pru p1 peach major allergen and are possibly functionally related. Plant Physiology 150: 1235-1247



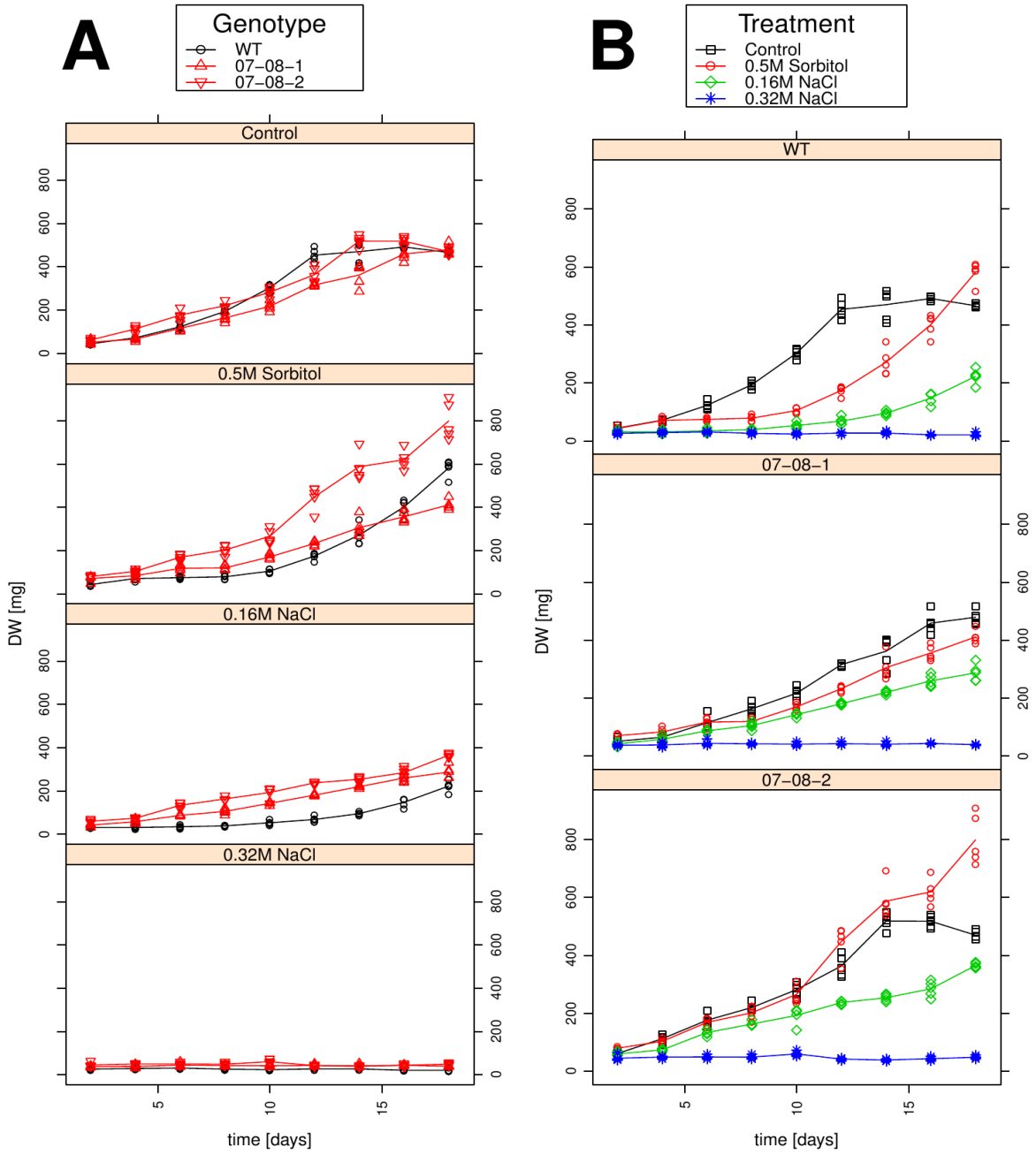
Chapter IV, figure 1. Time course of *pr-10a* gene-expression fold changes of cells treated with medium supplemented with (A) 0.5M sorbitol, (B) 0.16M NaCl and (C) 0.32M NaCl. Measurements for all five independent samples per time point and treatment group are displayed as symbols (circles and triangles) indicating individual measurements and lines giving the averages.



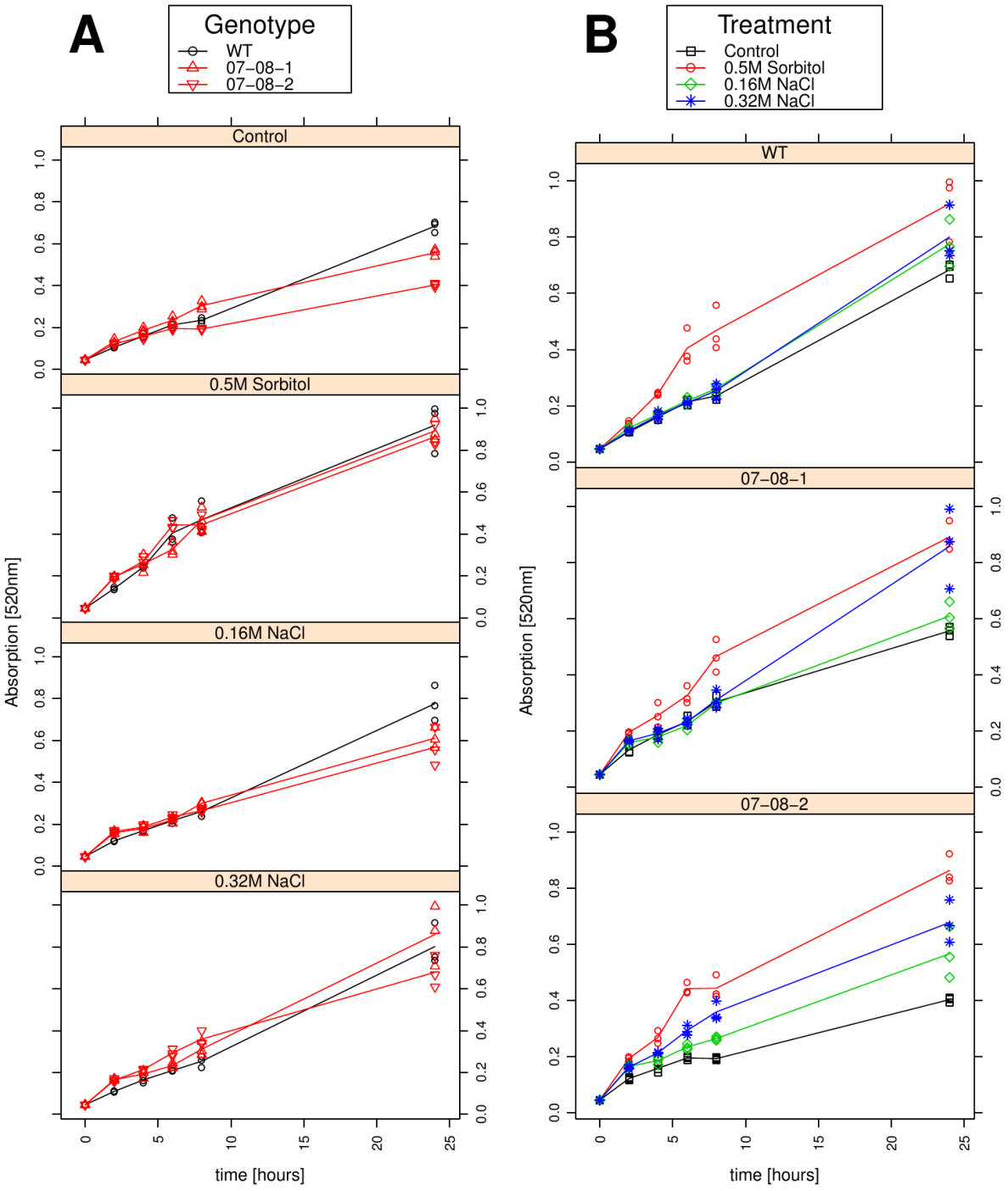
Chapter IV, figure 2. Time course of *sebf* (left column) and *pti4* (right column) gene-expression fold changes of cells osmotically stressed as indicated. Measurements for all five independent samples per time point and treatment group are displayed as symbols (circles and triangles) indicating individual measurements and lines giving their averages.



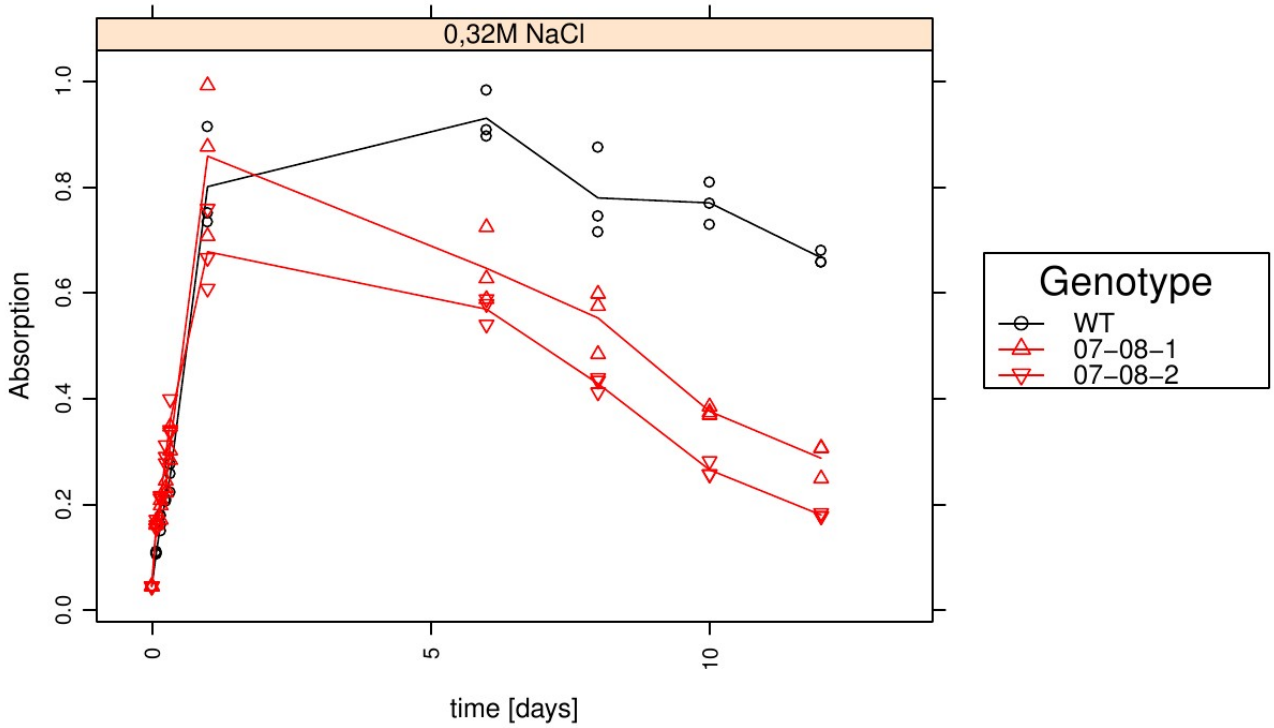
Chapter IV, figure 3. Time course of *luc* gene-expression fold changes in the transgenic cell lines under osmotic regimes as indicated. Measurements for all five independent samples per time point and treatment group are displayed as symbols (circles, rhombs and asterisks) indicating individual measurements per treatment and lines giving their averages.



Chapter IV, figure 4. Dry-weight accumulation of wild-type and transgenic cultures under control conditions and osmotic regimes as indicated. The dry-weight measures for all five Erlenmeyer flasks harvested every second day from wild-type and transgenic cultures are indicated. Measurements for all five independent samples per time point and treatment group are displayed as symbols (circles and triangles) indicating individual measurements and lines giving the averages. (A) Data are assigned to colours and symbols according to the genotypes. (B) Data are assigned to colours and symbols according to the treatments.



Chapter IV, figure 5. Respiration kinetics of wild-type (WT) and transgenic cultures under control conditions and osmotic regimes as indicated Measurements for all three independent samples per time point and treatment group are displayed as symbols (circles and triangles) indicating individual measurements and lines giving the averages. (A) Data are assigned to colours and symbols according to the different treatments. (B) Data are assigned to colours and symbols according to the treatments.



Chapter IV, figure 6. Respiration kinetics from wild-type (WT) and transgenic cultures under 0.32M NaCl treatment in liquid culture. The individual absorption measurements for all three Erlenmeyer flasks per genotype and time point (circles and triangles) and the averages (lines) are indicated.

Chapter IV, appendix

Southern-blot analysis

Genomic DNA was isolated from plant cells using the CTAB method of Doyle and Doyle (1990). For Southern blot analysis, 100µg of genomic DNA isolated from *Solanum tuberosum* cv. Désirée (WT) and the transgenic cell suspensions (07-08-1 and 07-08-2) and the plasmid *pGII 0229 MAS sth2 cp148 luc* as positive control were digested overnight with Fast Digest enzyme XbaI (Fermentas), electrophoretically separated on a 0.8% (w/v) agarose gel and subsequently transferred to positively charged nylon membranes (Roche) by the capillary transfer method. Primers used for *pr-10a* probe and *luc* probe were described in El-Banna et al. (2010). For DIG labelling a Roche Digoxigenin PCR probe syntheses kit (Roche Diagnostics GmbH, Mannheim, Germany) was used according to the manufacturer's instruction. Hybridization was carried out following standard procedures (Sambrook and Russell, 2001). A DIG luminescence detection kit (Roche CDP-Star, Roche Diagnostics GmbH, Mannheim, Germany) and a Fuji LAS 3000 imager were used for Southern Blot visualization.

Stable integration of the T-DNA was confirmed by Southern blot analysis. To discriminate between recombinant and endogenous *pr-10a* genes, hybridization of membranes was performed with a *pr-10a* probe and a *luc* probe (Figure 1). Hybridization signals in the wild type gDNA, hybridized with a *pr-10a* probe, originate only from the endogenous *pr-10a* gene, whereas hybridization signals in transgenic lines indicate integration profiles different from the wild type originating from a second endogenous *pr-10a*-gene and the recombinant gene(s). Apart from bands for the endogenous gene, for 07-08-1 one and for 07-08-2 two different loci for the recombinant *pr-10a* gene were detected. Interestingly, no common insertion loci for the two transgenic cell lines were obtained.

It has to be pointed out, that a cell culture represents a population of cells, which may consist of differently transformed cells. Therefore it is not possible to decide, whether the insertions found represent a unique cell type with two inserted gene copies or two different cell types, each containing one gene copy.

Magnitude of *pr-10a* expression induction is sensitive to both the chemical compound and the strength of osmotic pressure

Figure 2 displays the time course of relative expression levels of *pr-10a* (normalized against *18S* (Pfaffl et al. 2001)). Considering smaller ΔCt values indicate higher expression levels, both transgenic cell lines show substantially higher *pr-10a* expression levels from the beginning on.

All genotypes react on the mild 0.16M NaCl and the 0.5M sorbitol treatments with similar rates of expression induction. Due to the higher generally higher expression level of the transgenic cell lines, they perform parallely to the wild type cells throughout the observed period of time (Figure 2). However, the 0.32M NaCl treatment entails such a severe *pr-10a* expression induction in the wild type cells, that they reach the also heavily induced level of the transgenic cell lines (Figure 2).

Figure 3 and 4 display the time course of relative expression levels of *sebf* and *pti4* respectively (normalized against *18S* (Pfaffl et al. 2001). Considering smaller ΔCt values indicate higher expression levels, for *sebf* no explicit difference between wild type and transgenic cell lines can be determined throughout the different treatments (Figure 3). However, it could argue, that genotype 07-08-1 remains lower expression levels under control and 0.5M sorbitol treatment, while in both salt regimes the transgenic cell lines show higher gene expression levels throughout the observed period of time. But the detected differences in ΔCt values remain small.

For *pti4* ΔCt values did not indicate any substantial differences in gene expression level between the genotypes. However, the severe 0.32M NaCl treatment showed slight increase of gene expression levels for all three genotypes (Figure 4).

Growth behaviour is influenced by both the chemical compound and the strength of the osmotic pressure

For evaluation of growth curve results, a cell means model was set up describing the dry weight in dependency of the treatment (dry weight ~ treatment), whereas the treatment levels specify the combination of the genotype, the applied medium and the time of measurement, resulting in 108 levels.

For comparisons of experimental group means simultaneous multiple-comparison procedures according to Hothorn et al. (2008) and Schaarschmidt and Vaas (2009) were performed with R (R Development Core Team 2010) using the above-mentioned linear cell-means models.

Figure 5shows the results for a user defined set of mean comparisons revealing that both the type of chemical compound and its strength influence the growth.

Validation of TTC-based respiration curves

As recommended by Duncan and Widholm (1990), tests on background absorption, due to e. g. extractable pigments, and analysis of the resulting absorption spectra were performed. Therefore wild type cells were treated as intended but without Biolog Dye A. The absorption spectra of their ethanol extracts were measured and compared to spectra from samples containing middle and high amounts of formazan after 8h and 24h incubation in control medium supplemented with Biolog Dye A (Figure 6).

The ethanol extracts from treated cells without Biolog Dye A supplement did not result in accountable absorption values (Figure 6), indicating no disturbing cell compounds could be extracted, irrespectively of the applied treatment. Cells provided with Biolog Dye A produced the red formazan. Its spectrum exhibited a narrow peak with maximum at 520nm throughout different concentrations (red and green line). Thus 520nm was chosen for all absorption measurements determining formazan amounts.

Further possible interactions between the used dye, the applied media and compounds of deceased cells were investigated. Therefore cell material was inactivated by freezing and thawing before formazan formation in the different media supplemented with Biolog Dye A (Figure 7) was determined by measurement of the absorption at 520nm after 24h incubation time.

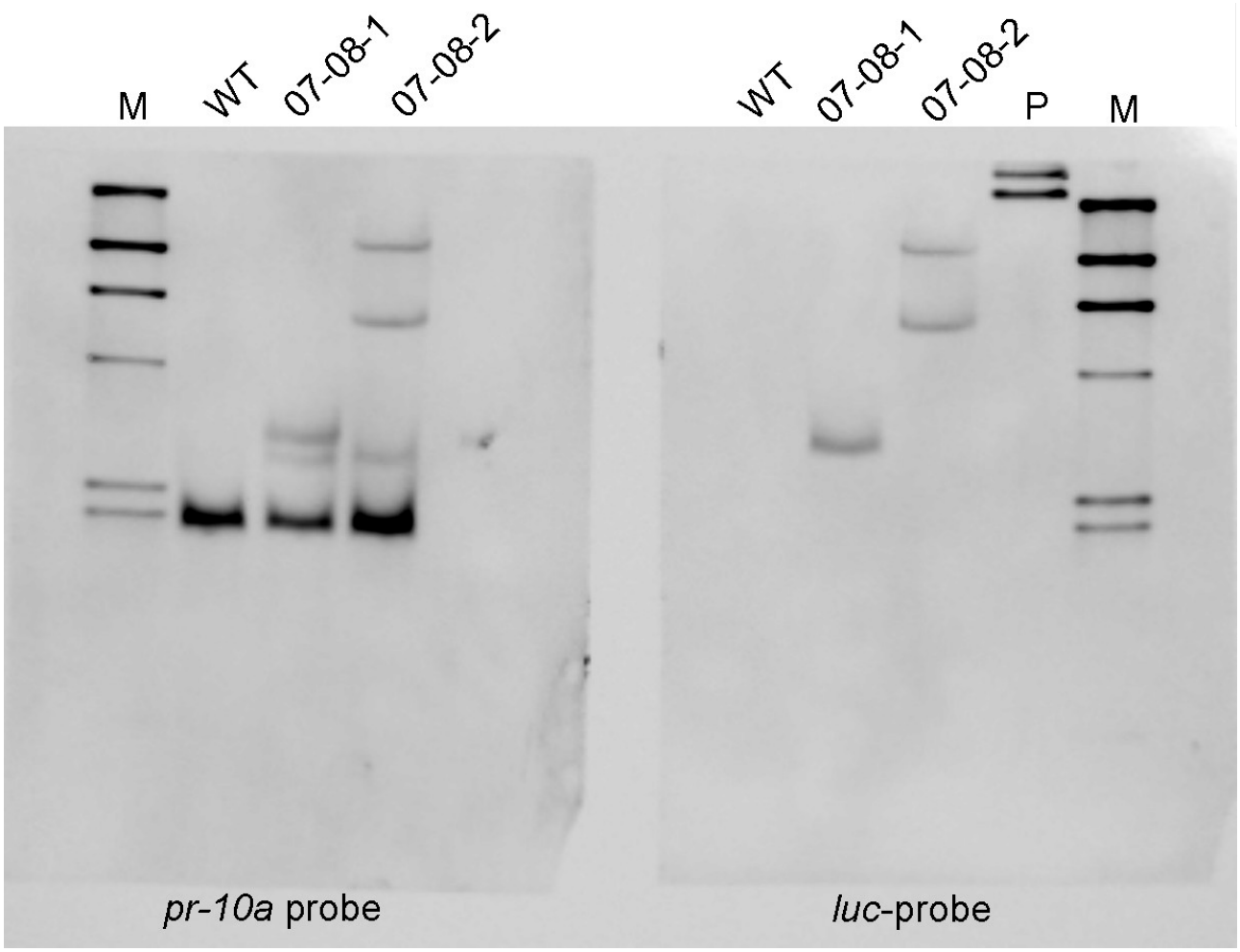
Neither interaction between the dye and the applied media itself, nor false positive formazan formation in the here used *Solanum tuberosum* cv. Desirée cell cultures was observed after inactivated by freezing and thawing (Figure 7 A and B). For comparison, absorption of ethanolic extracts from live WT cells under 0.5M sorbitol treatment supplemented with Biolog Dye A after 2 and 24 hours (very right bars) are given.

Respiration behaviour is influenced by both the chemical compound and the strength of osmotic pressure

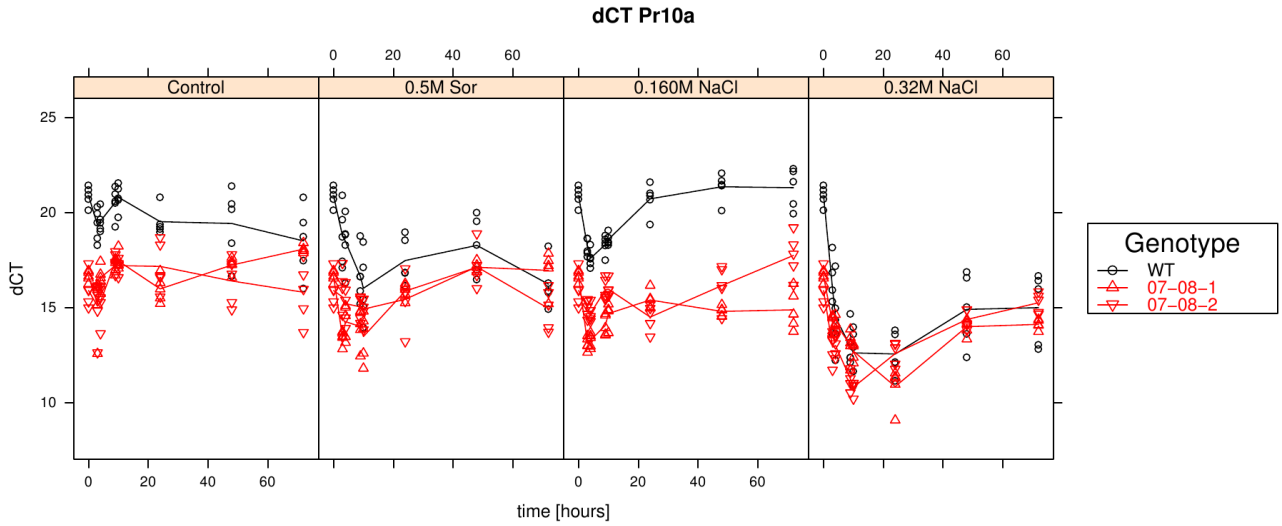
For evaluation of respiration curve results, a cell means model was set up, describing the absorption in dependency of the treatments (absorption ~ treatment), whereas the resulting 72 treatment levels consist specify the combination of the genotype, the applied medium and the time of measurement.

For comparisons of experimental group means simultaneous multiple-comparison procedures according to Hothorn et al. (2008) and Schaarschmidt and Vaas (2009) were performed with R (R Development Core Team 2010) using the above-mentioned linear cell-means models.

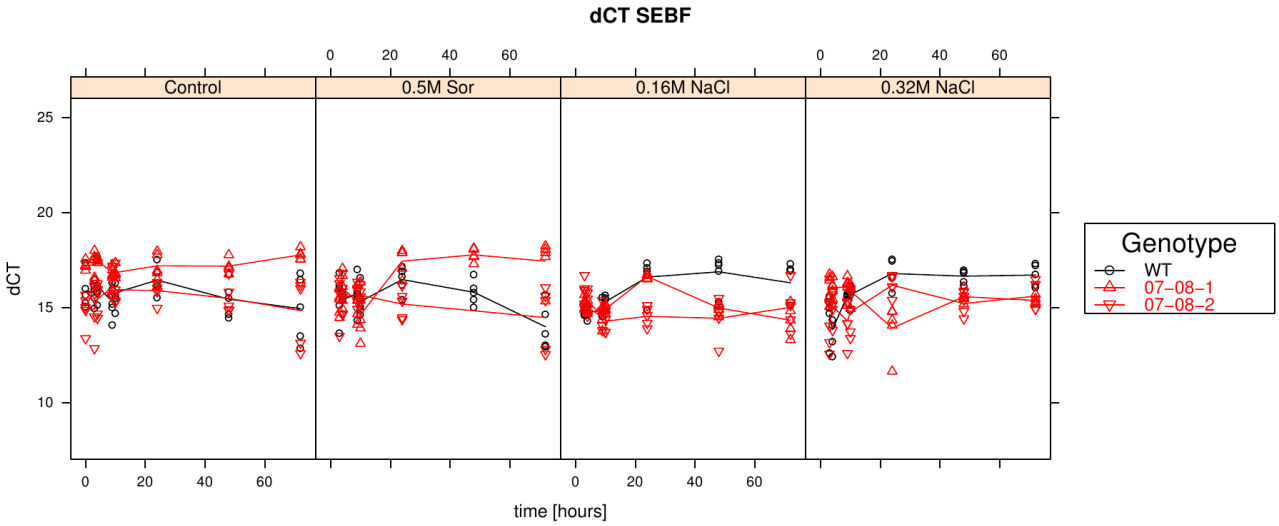
Figure 5 shows the results for a user defined set of mean comparisons revealing that both the type of chemical compound and its strength influence the respiration.



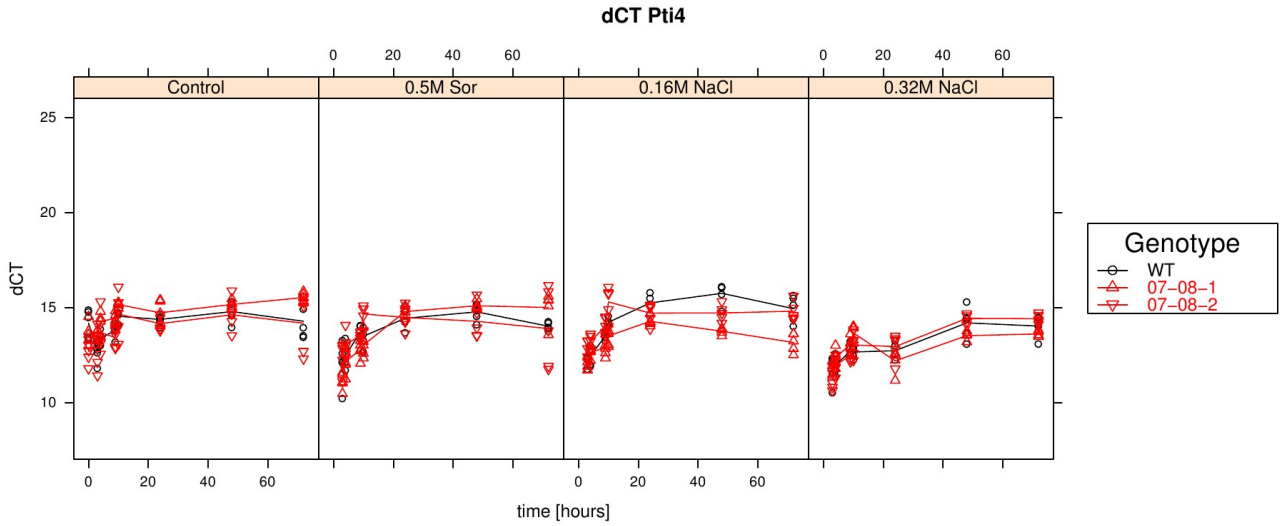
Appendix to chapter IV, figure 1. Southern blot analysis of DNA from transgenic and wild type potato suspension cells. DNA was digested with XbaI. Digoxigenin labelled hybridization probes for the *pr-10a* as well as for luc gene have been used. M: digoxigenin-labelled DNA Molecular Weight Marker II (Roche); 07-08-1 and 07-08-2: transgenic clones of *Solanum tuberosum* cv Désirée; WT: wild type clone; P: plasmid DNA.



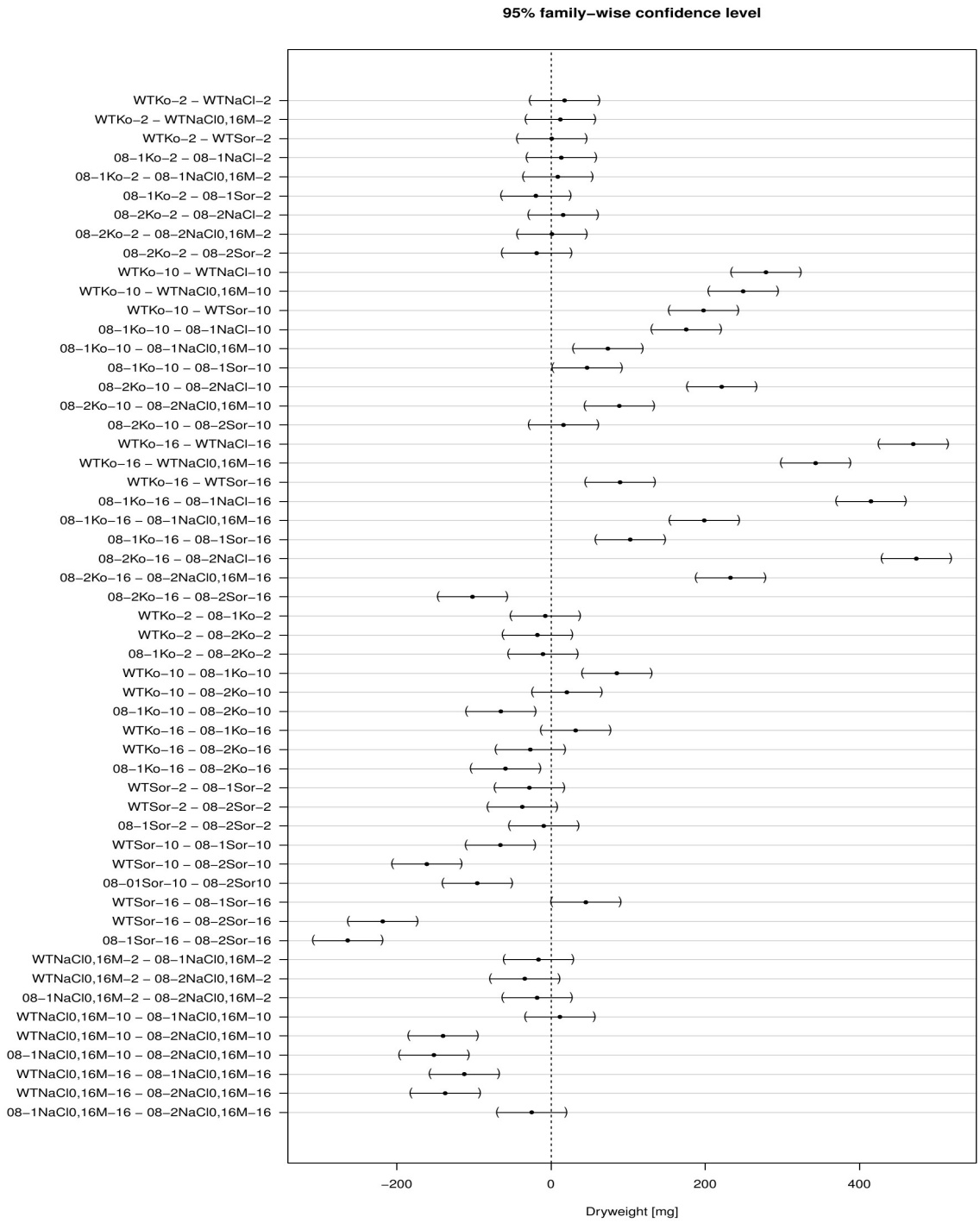
Appendix to chapter IV, figure 2. Results of relative *pr-10a* expression analysis of cells treated with differently supplemented medium (Control, 0.5M sorbitol, 0.32M NaCl, 0.16M NaCl). Given are the expressions levels standardized to their *18S* rRNA threshold cycle (ΔCt). WT is coded with black lines and circles, GMO with red lines and triangles. Note that smaller ΔCt values indicate a stronger expression (Pfaffl 2001).



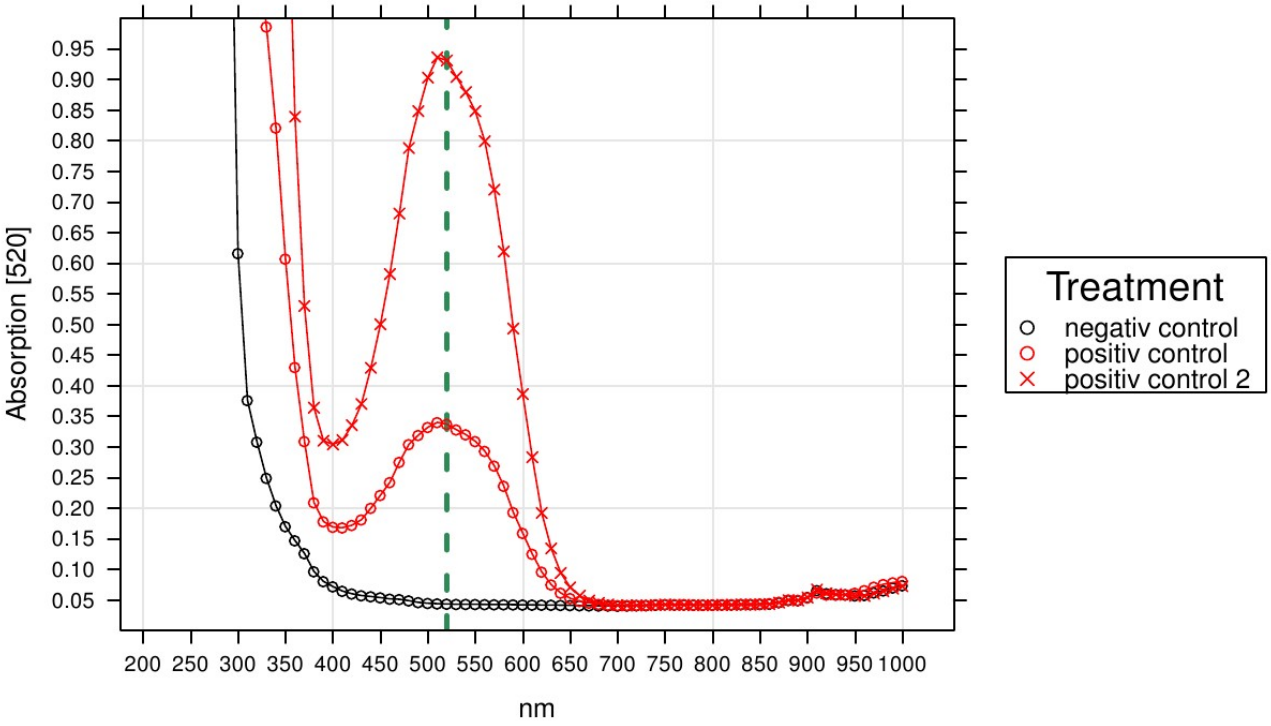
Appendix to chapter IV, figure 3. Results of relative *sebf* expression analysis of cells treated with differently supplemented medium (Control, 0.5M sorbitol, 0.32M NaCl, 0.16M NaCl). Given are the expressions levels standardized to their *18S* rRNA threshold cycle (ΔCt). WT is coded with black lines and circles, GMO with red lines and triangles. Note that smaller ΔCt values indicate a stronger expression (Pfaffl 2001).



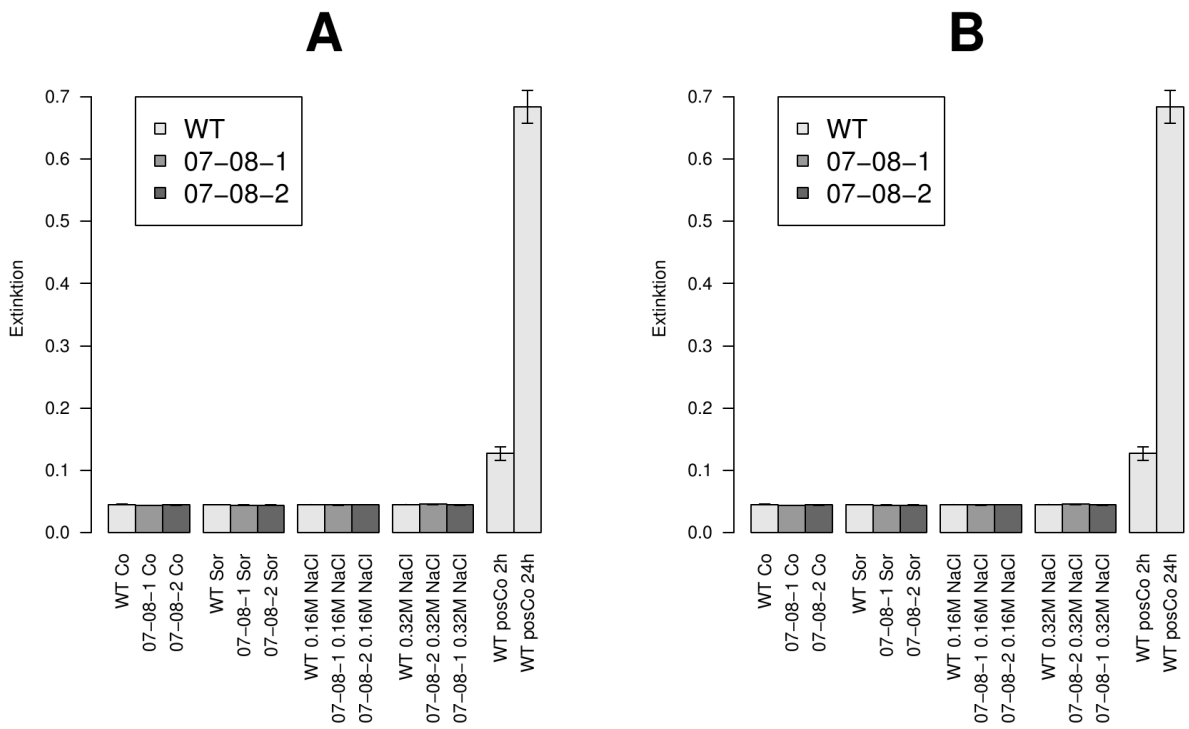
Appendix to chapter IV, figure 4. Results of relative *pti4* expression analysis of cells treated with differently supplemented medium (Control, 0.5M sorbitol, 0.32M NaCl, 0.16M NaCl). Given are the expressions levels standardized to their *18S* rRNA threshold cycle (ΔCt). WT is coded with black lines and circles, GMO with red lines and triangles. Note that smaller ΔCt values indicate a stronger expression (Pfaffl 2001).



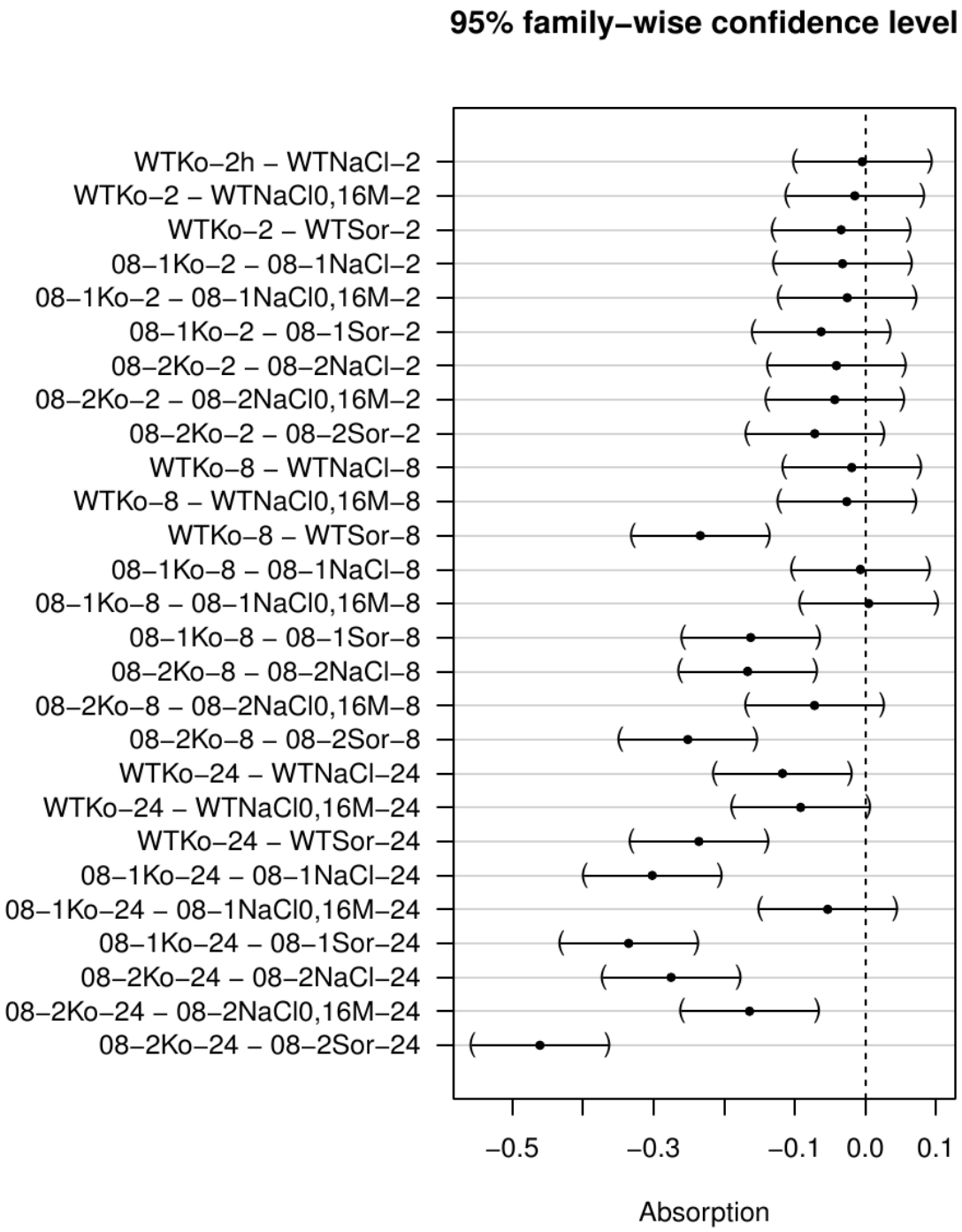
Appendix to chapter IV, figure 5. User-defined set of differences of means calculated using simultaneous multiple comparisons result Confidence intervals (95%) for dry weight accumulation under different treatments and after different incubation times. Tukey-type contrasts for each genotype separately was calculated for 2, 10 and 16 days of incubation time. Additionally means of the different genotypes under corresponding conditions were compared for these incubation times. A 95% CI bar crossing the dotted vertical zero line indicates the absence of a significant difference between the considered samples.



Appendix to chapter IV, figure 6. Absorption spectra of ethanol extracts of wild type cells without Biolog Dye A (black circles), ethanol extracts of wild type cells treated with 4X medium after eight (red circles) and 24 hours (red crosses). The green dashed line indicates the absorption maximum at 520nm.



Appendix to chapter IV, figure 7. Absorption ethanol extraction of differently treated cells after 24h. Given are mean and standard deviation from three independent replicates. WT indicated by light grey bars, 07-08-1 indicated by grey bars, 07-08-2 by dark grey bars, respectively. (A) Absorption at 520nm of ethanol extracts from differentially treated cells without Biolog Dye A supplement. For Comparison absorption of ethanol extracts from live WT cells under 0.5M sorbitol treatment supplemented with Biolog Dye A after 2 and 24 hours (very right bars). (B) Absorption at 520nm of ethanol extracts from inactivated cells supplemented with Biolog Dye A. For comparison absorption of ethanol extracts from live WT cells under 0.5M sorbitol treatment supplemented with Biolog Dye A after 2 and 24 hours (very right bars) are given.



Appendix to chapter IV, figure 8. User defined set of differences of means calculated using simultaneous multiple comparisons result Confidence intervals (95%) for absorption at 520nm of formazan in ethanol extracts from cells under different treatments and after different incubation times. A 95% CI bar crossing the dotted vertical zero line indicates the absence of a significant difference between the considered samples.

V. Exploiting the wealth of Phenotype MicroArray data: efficacious visualization of and robust parameter estimation from respiration kinetics

Abstract

The Phenotype MicroArray (OmniLog® PM) system is able to simultaneously capture a large number of phenotypes by recording an organism's respiration over time on distinct substrates. This technique targets the object of natural selection itself, the phenotype, whereas previously addressed '-omics' techniques merely study components which finally contribute to it. The recording of respiration over time, however, adds a longitudinal dimension to the data. To optimally exploit this information, it must be extracted from the shapes of the recorded curves and displayed in analogy to conventional growth curves.

The free software environment R was explored for both visualizing and fitting of PM respiration curves. Approaches using either a model fit (and commonly applied growth models) or a smoothing spline were evaluated. Their robustness in inferring curve parameters and confidence intervals was compared to the native OmniLog® PM analysis software. We consider the post-processing of the estimated parameters, the optimal classification of curve shapes and the detection of significant differences between them, as well as practically relevant questions such as detecting the impact of cultivation times and the minimum required number of experimental repeats.

We provide a comprehensive framework for data visualization and parameter estimation according to user choices. A flexible graphical representation strategy for displaying the results is proposed, including 95% confidence intervals for the estimated parameters are facilitated. The spline approach is more robust against irregular curve shapes than fitting any of the considered models or using the native PM software for calculating both point estimates and confidence intervals. These can serve as a starting point for the automated post-processing of PM data, providing much more information than the strict dichotomization into positive and negative reactions. Our results form the basis for a forthcoming freely available R package for the analysis of PM data.

Introduction

The so called '-omics' techniques yielded tremendous insights in the biology of cellular organisms. They address different steps in the information transfer from coding DNA (genomics) *via* RNA (transcriptomics) to the proteins (proteomics and interactomics) to finally yield the cellular metabolites (metabolomics and fluxomics) (Buchanan et al. 2009; Vilchez-Vargas et al. 2010; Zhang et al. 2010;). Other '-omics' techniques are MicroRNomics, probiogenomics, lipidomics and fluxomics (Kandpal et al. 2009; Ventura et al. 2009; Wenk 2010; Feng et al. 2011). Their unifying theme is the study of the cellular totality of the organisms of interest to obtain a systematic insight into basic biology (Bujara and Panke 2010; Ruppin et al. 2010) and to reconstruct complex metabolic networks and flow-charts of fluxes (Nogales et al. 2008; Le Cao et al. 2009; Medina et al. 2010; Mithani et al 2011). The data flood to be processed is enormous, depending on the experimental setup.

A major biological feature, the phenotype, was until recently not accessible with high-throughput techniques. This is unfortunate, as it is the phenotype which is the object of selection and, hence, is the level at which evolutionary directions are governed (Mayr 1997). All previously addressed '-omics' techniques merely study components which finally contribute to the phenotype (Papin et al. 2004).

In microbiology, a simple way to assess the phenotype is to characterize an organism's replication behaviour under specific conditions (Neysens et al. 2003; Sikorski et al. 2007) by analysing the shape of the growth curve during the commonly known growth phases. The length of the lag phase reveals how fast and well the organism acclimates to a specific environmental condition, while the period of cell replication, the log phase, and the stationary phase (when growth comes to an end) indicate the particular way the growth is achieved (Brisbin et al. 1987). Unfortunately, manually recording growth curves is an extremely time- and cost-intensive work.

The Phenotype MicroArray (PM) system appears to close the gap of capturing a large number of phenotypes in high-throughput systems. In this approach, a physiological reaction producing NADH engenders a redox potential and flow of electrons to reduce a tetrazolium dye (Bochner and Savageau 1977) such as tetrazolium violet, thereby producing purple colour. The more rapid this metabolic flow, i.e. cellular respiration, the faster the formation of purple colour (Bochner et al. 2001; Bochner 2009). The OmniLog® PM system records the colour change every 15 minutes in an automated setting under up to 2000 distinct physiological challenges, such as the metabolism of single carbon sources, metabolism under

varying osmolyte concentrations, and response to varying growth-inhibitory substances (Bochner et al. 2001; Bochner 2009). The challenges can be further augmented by modifying environmental conditions such as the temperature and the composition of the gaseous phase.

In common '-omics' techniques, the recorded value is a mostly qualitative information on the difference between two experiments, usually obtained from measurements at a single time point, which is often an endpoint (Lay et al. 2006). In contrast, the PM respiration kinetics add a longitudinal dimension. This higher level of PM data complexity contains additional valuable biological information coded in the shape characteristics of the recorded curves in analogy to conventional growth curves as introduced above (Brisbin et al. 1987). These curve features can, in principle, unravel fundamental differences or similarities in the respiration behavior of distinct organisms, which cannot be identified by endpoint measurements alone.

This wealth of data was till now hardly exploited, as the kinetics were usually only qualitatively assessed (Haack et al. 1995; Homann et al. 2005; Sabet et al. 2009; Di Cagno et al. 2010; Xue et al. 2011). The mere classification into a positive or negative reaction to an environmental challenge appeared to be sufficient, whereas the kinetic information itself was neglected. Also, the application of PM in functional genomics, as, e.g., for improving genome annotation (Reed et al. 2006) and assessing gene function using knock-out techniques, exploits only presence/absence calls (Puchalka et al. 2008). Nevertheless, already these early studies exhibited the complexity of the situation by, in the light of current knowledge, completely unexpected or even incomprehensible results (Zhou et al. 2003). Even though the need for a more sophisticated strategy for data analysis was emphasized long ago (Warringer and Blomberg 2003), only data recording could be accelerated until now. Although first attempts to establish web-based data storage and analysis infrastructures were already made (Chang et al. 2011), an efficient bioinformatic evaluation tool that includes all steps of longitudinal-data analysis, or even a methodical collection analogous to BIOCONDUCTOR for conventional microarray data (Gentleman et al. 2004), is still unavailable.

The native OmniLog® PM software (Biolog 2009) displays the PM measurements only according to the 8×12 wells plate layout and provides only limited functionality for the visual comparison of at most two curves. The PM software includes a parametric analysis, which calculates parameters describing a curve's kinetic shape but disregards modeling or curve-fitting approaches and does not provide confidence intervals (CIs), even though it is well known that these can be used to examine statistically detectable differences (Schenker and Gentleman 2001; Cumming et al. 2007). Third-party tools include data visualization

(Jacobsen et al. 2007), but to the best of our knowledge are not publicly available. Some simple but effective approaches to data analysis using summary statistics of growth curves (Hackett and Griffiths 1997) or hypothesis-testing frameworks (Sturino et al. 2010) were also published, but these approaches reduce the information content of each curve to one or a few single values and use these to determine respiration differences on the various substrates without considering the curve shapes.

The development of statistical methods for the analysis of longitudinal data started with the pioneering work of Laird and Ware (1982) which discussed a general family of models including growth models and repeated-measures models as special cases. Studies on nonlinear and linear mixed-effects models, the integration of splines, random coefficients and variance modeling into a flexible analysis approach based on linear mixed models followed this seminal work (Härdle 1992; Laird et al. 1992; Eilers and Marx 1996; Birch 1999; Verbyla 1999; Serroyen et al. 2009). Highly elaborated tools for the evaluation of longitudinal data are implemented in statistical software such as the packages *drc* (Ritz and Streibig 2005) and *grofit* (Kahm et al. 2010) in R, *PROC MIXED* in SAS (Singer 1998), *xtmixed* in Stata (Rabe-Hesketz and Skrondal 2008), and *MIXED* in SPSS (Peugh and Enders 2005). Also, many mathematical models describing growth behaviour have been developed (Zwietering et al. 1990; Neysens et al. 2003; Mitchell et al. 2004).

Most empirical equations such as the logistic law (Feller 1940) or Richards curves (Richards 1959) fit well onto growth data *via* plain non-linear regression if the growth follows the typical sigmoid shape, but mathematical simplicity also plays a key role (Mitchell et al. 2004). Hence, the application of these models to even slightly non-typical growth behavior (e.g., the simple violation of the assumption of symmetry around inflection) can lead to systematic errors (Gottschalk and Dunn 2005) and potentially to biologically unreasonable results (see below). To overcome this problem, the best-fitting model can be detected using the Akaike information criterion (AIC), which balances between fit and model simplicity (Akaike 1973; Akaike 1983). Unfortunately, general guidelines for the selection of the types of models to test are unavailable. Spline smoothers (Eilers and Marx 1996; Serroyen et al. 2009) are an alternative to describe growth or respiration behaviour, particularly if violations of model assumptions are both common and also reveal biologically important information.

Here we explored the free software environment R (R Development Core Team 2010) for both data visualization and fitting of growth curves for the comparative analysis of PM data. R is one of the most widely used solutions for statistical computing, featuring powerful

interactive data exploration as well as programming tools and numerous add-on packages. We first assessed the suitability of the *lattice* package (Sarkar 2008) for (re-)implementing and comparing previously published (Jacobsen et al. 2007) and alternative strategies for raw data visualization of 9792 bacterial respiration curves. Second, we examined which kinds of divergences from typical sigmoid growth curves occur, which kinds of artefacts might affect the reproducibility of the results and, hence, which basic quality-control measures are necessary and can be performed using the here presented software tools. Third, following the model-fitting approach of (Fodor et al. 2005) we assessed the *grofit* package (Kahm et al. 2010) for automatically conducting model fits as well as model-free fits using spline smoothers. The robustness of both approaches when inferring curve parameters (and their CIs) from PM data was compared with the current implementation in the native OmniLog® PM analysis software (Biolog 2009) and the specific merits and deficiencies of either method were determined. Fourth, we applied the tools to research questions relevant for establishing settings for OmniLog® PM production runs, illustrating how the experimenter can detect significance and magnitude of differences between the considered curve parameters to ensure reproducibility of the results in accordance with predefined quality standards (Gardner and Altmann 1986). Finally, as another example for the post-processing of the inferred parameters, we classified the curves into characteristic shapes. In contrast to the typical dichotomization of PM curves into occurrence of respiration and lack thereof (Xue et al. 2011), we here inferred curve archetypes (Eugster et al. 2009) to explicitly address the question of how many, and which, classes of curve shapes optimally represent the data.

Our results enable us to provide recommendations for software solutions for exploiting multiple respiration kinetics from automated systems such as PM. We believe that the availability of convenient and robust data exploration techniques *via* freely available software such as R will allow users of the PM technology to conduct in-depth data analyses that go significantly beyond the consideration of mere endpoint measurements and presence/absence calls.

Material and Methods

Organisms studied and PM measurements conducted

The first dataset comprised two strains of two species of bacteria (*Escherichia coli* DSM 30083^{T}, *E. coli* DSM 18039 = K12, *Pseudomonas aeruginosa* DSM 1707 and *P. aeruginosa* 429SC). *P. aeruginosa* DSM 1707 was grown on M1 agar (5 g/l peptone, 3 g/l meat extract

(Oxoid), 15 g/l agar); all other strains were grown on LB medium (lysogeny broth; 10 g/l peptone, 5 g/l yeast extract, 10 g/l NaCl, 15 g/l agar) for nearly 24 h and subsequently measured on GEN III MicroPlates™ (AES Chemunex BLG 1030) in the PM modus over 91 h. Each strain was measured in ten technical replicates. To ensure that all ten replicate plates were inoculated with cells of identical physiological conditions, the desired cell concentration was adjusted in a pool of ten vials of GEN III inoculation medium A (AES Chemunex BLG 72401) which was then simultaneously inoculated into ten GEN III plates. The second dataset followed the same design, but was collected two weeks later, thus representing a biological repetition. The two datasets thus comprised a total of four strains × two biological replicates × ten technical replicates × 96 substrates, hence 7680 individual curves.

To additionally investigate the impact of the age of cultures on the technical and biological reproducibility, the third dataset focused on a single strain only, *E. coli* DSM 18039 = K12, which was grown on solid LB medium for 16.75 h (t_1), 18.00 h (t_2), 19.33 (t_3), 20.50 (t_4), 21.92 (t_5), 23.25 h (t_6), 24.5 h (t_7), 25.58 h (t_8) or 40.33 h (t_9), respectively, and subsequently measured on GEN III MicroPlates™ in the PM modus over 91 hours. For each growth duration age four technical replicates were performed except for t_6, which was repeated only twice. Dataset 3 thus comprised one strain × eight growth durations × four technical replications × 96 substrates plus (t_9) one strain × one growth duration × two technical replicates × 96 substrates, hence 3072 + 192 = 3264 individual curves.

Visualization of PM raw data

As the functionality of the native OmniLog® PM software (Biolog 2009) is specialized on only few functions (see above) we first used the add-on package *lattice* (Sarkar 2008) for R (R Development Core Team 2010) to visualize the PM curves as heat maps using the function *levelplot()*, equivalent to a re-implementation of the approach of (Jacobsen et al. 2007). We then applied *lattice* to explore alternative visualization strategies using trellis graphics, which arrange graphics in a regular grid-like structure. Large and complex structured datasets can be regularly subdivided according to variables from the chosen experimental design, and in each panel one subset can be graphed, finally providing coordinated, high-dimensional views (Becker et al. 1996). As curves are the most comprehensive display of kinetics, we used the high-level function *xyplot()*, which can plot curves in any requested sub-division, combination and constellation. We examined which display method provided the most natural way to assess data quality and data integrity. The main potential artifacts, the range of potential curve

shapes and other issues potentially affecting measurement reproducibility were identified during this step by visual inspection of all curves.

Parameter estimation from respiration curves

For the description of functional dependencies of two measured variables a mathematical function can be fitted onto the data. In general, such a fit aims at minimizing the distances between the raw data points and the values predicted by the function. The choice of a type of function is usually motivated by some basic assumption about the underlying system. The selection of a function is an interpreting activity and a crucial step in the analysis (Hütt 2001). Alternatively, the dependency between two measured variables can be described by smoothing splines. Those splines can be thought of as a concatenation of cubic polynomial segments that are joined together at their ends or knots (Wold 1974). Their unique property as an empirical function is that they can represent any variation in curve shape.

The parametric analysis method of the native OmniLog® PM software (Biolog 2009) only crudely accesses possible differences in curve shapes, because it uses only few data points from the curve for the computation of curve kinetic parameter values (see p. 38 in chapter 5 of the OmniLog® user guide (Biolog 2009)). The maximum height ("MaxHeight") is given as the 10^{th} percentile highest value among all values over all time points, and the minimum height ("MinHeight") is calculated as the 12^{th} smallest value among the first 48 reads over all time points. The length of the lag phase is calculated from the raw data using the formula "MidTime - (MidHeight – MinHeight) / Slope" (Biolog 2009), while "MidTime" is described as the first time a value exceeds MidHeight. "MidHeight" is defined as the value midway between MinHeight and MaxHeight. The Slope is calculated as "sum of rises over run between 15% Time and MidTime –1 and rises over run between MidTime +1 and 85% Time" divided by "85% Time minus 15% Time" from the raw data (Biolog 2009). Here, "x% Time" is defined as the first time a value exceeds the value x% of the way between MinHeight and MaxHeight. The calculation of the area under the curve (AUC) is described as "the sum of all OmniLog values over all time points (area under the curve)" (Biolog 2009), which treats the colour changes between time points as a step function. Also, native OmniLog® PM software only provides point estimates but not CIs, which are important for statistical evaluations (Schenker and Gentleman 2001; Cumming et al. 2007). Hence, the software cannot be used to investigate whether two quantitatively similar curves differ in a statistically detectable way.

In contrast, the basic part of R's add-on package *grofit* (Kahm et al. 2010) provides a framework for parameter estimation using model fitting and model-free spline fitting separately and also allows the statistical assessment of the curves using CIs. The model-based approach fits each predetermined model by a non-linear least-squares regression. The Akaike Information Criterion is used to select a best model. The spline fitting approach is based on a cubic smoothed spline and follows the framework implemented in the R function *smooth.spline()*. We here applied the default smoothing parameter. The package *grofit* (Kahm et al. 2010) was originally built to derive dose-response curves and calculate descriptive pharmacological or toxicological values. The here used intermediate output contains estimates for the following parameters: the length of the lag phase λ, the growth (here: increase in respiration) rate μ (corresponding to "slope") and the maximum cell growth (here: respiration) A (corresponding to the maximum value recorded). As an additional descriptive parameter of cell growth (here: respiration), the area under the curve (AUC) is estimated via numerical integration (see Figure 2 in (Kahm et al. 2010) for details). In addition to the point estimates for the parameters from both model and spline, confidence limits can be calculated *via* bootstrapping, with 95% being the default value (Efron 1979). Significant differences can then be detected as non-overlapping CIs. In case of no overlap, the differences between the opposite limits of the considered CIs describe the smallest expectable mean difference.

We assessed in detail in how many (and which) cases a model fit was impossible using one of the default models: (i) logistic growth, (ii) Gompertz growth, (iii) modified Gompertz growth and (iv) Richards growth (Kahm et al. 2010). Particular emphasis was laid on biologically unreasonable parameter estimates as observed in preliminary experiments (data not shown) such as negative values for λ and estimates for the A exceeding 400 OmniLog® units (due to technical restrictions, the current version of the OmniLog® device yields measurements at most 400 OmniLog® units in height; Bochner, pers. comm.). To provide a rough estimate for the proportion of positive and negative reactions, we applied a (partially arbitrary) threshold of 100 OmniLog® units, i.e. larger estimates of A indicated positive reactions, other values indicated negative reactions (for a more advanced treatment see the inference of archetypes below). We also determined the correlations between all four parameters from the same curve fitting as well as between those from model fitting and spline fitting applied to the same raw data. Spearman's correlation index and Kendall's τ were compared, since the data are not necessarily normally distributed and the relationships not necessarily linear.

The accuracy with which the parameters (estimated using the three different approaches) fitted to the raw data was investigated for all types of observed curve shapes (see above) and

visualized using a set of individual curves representative for each shape. These exemplars could also be used to illustrate the difference in parameter estimation between model and spline fit and thus for the identification and explanation of the effect of difficult-to-fit curve shapes on parameter estimates. Moreover, they were used to determine the most useful way of displaying parameters estimated together with their CIs. The proposed methods here intentionally resign any multiplicity adjustment, because the analyses are expected to detect all interesting phenomena while it would be worse to miss some of them.

Detecting significant differences

Because there is no restriction on the type of sample to be analysed, the PM technique is capable of dealing with a rather unlimited amount of distinct experimental questions. That is, not only isolated strains or well-defined mutants are manageable, but also mixed or environmental samples are feasible (Spiegelman et al. 2005; Christian and Lind 2006). For most of them predictions about their behaviour are impossible, thus the experimenter needs to compare repeated measurements to be able to assess the range of variability in the specific sample, strain, etc. To demonstrate the value of CIs for data evaluation, we assessed scenarios where (i) curves differ significantly in general, (ii) replications differ significantly in some parameters but not in others, and (iii) differences between replications are not statistically detectable, as indicated by the 95% CIs. Such exemplars were also used to determine efficient ways to display these differences. As a laboratory example, we calculate 95% CIs from the third dataset to assess whether there was a significant impact of the age of the bacterial inoculation culture on technical and biological reproducibility. Apparently, if such a dependency was detected in a real-world dataset, the experimenter would need to more strictly standardize cultivation times prior to conducting PM measurements. Since up to now the *grofit* package is not intended for fitting a single model or spline on a set of several repetitions of a longitudinal data set, we present two alternative approaches for their comparison. First, we provide a graphical solution which yields preliminary insights into the overall behaviour of the considered groups and is based on mean parameter estimators and mean CIs calculated by averaging the corresponding values estimated from the individual curves. Second, as a somewhat more sophisticated approach, we provide a simultaneous multiple comparison procedure of means (Bretz et al. 2010). It provides test decisions using 95% CIs for the differences of parameter means according to a user-defined set of comparisons.

The behaviour of the negative controls (well A01) was examined more closely, particularly regarding the question whether it is valid to subtract these values from the measurements from all other wells before estimating curve parameters, a procedure which is sometimes recommended (Biolog 2009). Apparently, this strategy assumes a biologically sensible additivity between the negative control and respiration reactions caused by the substrates. Our results are presented in the appendix to this chapter.

Which, and how many, classes of curve shapes?

To explicitly infer the optimal number of distinct shape classes for classifying the curves, we applied the R package *archetypes* (Eugster et al. 2009) to the parameter estimates from the spline fits from the 1st and 2nd dataset. Archetypes are characteristic extreme types of combinations of multivariate observations found by minimizing a convex residual sum of squares (RSS) criterion; the implementation ensures that the real measurements can be represented by convex combinations of the archetypes. The algorithm alternates between finding the best set of coefficients for the given archetypes and finding the best archetypes for a given set of coefficients. The overall RSS is reduced successively because in each step several convex least-squares problems are solved. While the number of archetypes is predefined in each run, the algorithm can be restarted for a series of numbers of archetypes (we tested 1 to 10) and, according to the "elbow criterion", the optimal number is the one that resulted in the largest step towards a lower RSS compared to subsequent improvements. We used the *stepArchetypes()* function with five random starts per given number of archetypes. As some of the spline estimates for the parameter λ were outliers below zero (see below), its distribution was truncated (made symmetrical) by setting all values lower than the maximum times -1 to this value.

Results

Visualization of PM raw data

Using a subset of dataset 1 the visualization of PM curves as heat maps *via* the *lattice* function *levelplot()* is shown in Figure 1. The user was free to define any ordering of the lines in the columns, since the well position on the 8×12 GEN III MicroPlate™ is given on the y-axis and identification was easily possible. This also allowed the comparison of technical and/or biological replicates after an appropriate re-arrangement (data not shown). One advantage of this visualization technique was that numerous curves could be displayed in relatively small

space, when vertical lines representing the respiration curves were stacked (Figure 1). Further data quality assessment was feasible straightforwardly; for instance, deviations from the expected monotonic increase of the curve height could be identified and located (Figure 1).

In Figure 2 kinetic data are plotted according the original 8×12 well layout as superimposed curves by using the function *xyplot()*. Here, the user was also free to rearrange distinct numbers of curves in individual compositions of panels. Distinct organisms or treatments could be colour-coded (Figure 2) or alternatively indicated by distinct line types (not shown). Such a graphical display easily enabled to monitor the general performance of an organism and to simultaneously identify potential artifacts, such as individual replications that deviate in curve height or shape, or other irregularities such as deviations from the expected monotonic increase of the curves, particularly when superimposing the curves from distinct replicates of the same reaction in one panel.

We felt that data quality and integrity could be checked faster and more comprehensively using the second, curve-based visualization approach. The curve display gave a more intuitive and straightforward overview of the data, while simultaneously facilitating the development of an overall assessment of an organism's behaviour in the experiment. Moreover, colour codes for results from distinct organisms, replications and experiments enabled informative superimposed displays (Figure 2), which would be difficult when colour is used to indicate signal strength (Figure 1). For this reason, we used the visualization approach of Figure 2 to inspect the curves from all PM experiments. By this, we found various combinations of negative reactions, where (nearly) no color was formed in the wells, and positive reactions per organisms and/or experiment, particularly between the distinct biological replicates. It also turned out that a surprisingly large number of curves from positive reactions diverged from the typical sigmoid shape of growth curves (Figures. 2, 3). In most sets of technical and/or biological replicates which included such deviating curve shapes, these occurred in all of the respective replicates.

Parameter estimation from respiration curves

We estimated the parameters length of the lag phase λ, respiration rate μ (slope), maximum cell respiration A and the area under the curve (AUC) using both the model fitting and model-free spline fitting approach from the basic part of R's add-on package grofit (Kahm et al. 2010). Summary statistics from parameter estimation are shown in Table 1.

Depending on the specific dataset, between 1.4% and 6.4% of the curves could not be fitted by the modelling approach. Hence, for some experimental groups no parameter estimation was possible at all, resulting in one to twelve groups without parameter estimation depending on the dataset. In contrast, the spline resulted for all datasets, yielding parameter estimates for every group (Table 1). As mentioned above, biologically reasonable values for λ can, in principle, not be negative or exceed the last time point of measurement, whereas a reasonable A should be a positive value not greater than 400. Slightly negative values and those only slightly exceeding 400, however, can be judged as just negligibly inaccurate estimations of 0 and 400, respectively. The model-fitting approach resulted in negative estimates for λ in 23.7% to 36.6% of the groups and in A estimates exceeding 400 in 0.4% to 1.4% of the groups (yielding at least one uninterpretable parameter in between 25.4% and 39.5% of the groups).

In contrast, the spline fit yielded negative λ in only 14.4% to 28.8% of the groups; hence around 10% fewer groups with unreasonable λ point estimators. Not a single group was found with an estimate for A exceeding 400 (Table 1). Accordingly, uninterpretable values for one parameter (λ), if any, did not result in uninterpretable values for others (A). For those spline estimates with A > 100 (approximately representing positive reactions), only between 4.0% and 4.9% of the lambda values were negative, and only slightly so (mean between -3.3 and 3.0 h). The vast majority of negative λ occurred for A < 100 (approximately representing negative reactions). In the cases of datasets 2 (mean -6.8) and 3 (mean -12.1), these values were also only slightly negative. Only in the case of dataset 1, additionally a number of extremely low λ estimates were encountered (mean -207.6). There was little difference between model fitting and spline fitting regarding the estimated proportion of negative reactions (Table 1).

Kendall and Spearman correlations between the parameters describing the curves are listed in Table 2. In the model-fitting framework the correlation between λ and the other parameters was quite low. Also, μ was moderately correlated with A but more strongly with AUC (0.732/0.712). The correlation between A and AUC was a bit lower (0.700/0.522). Within the parameters from the spline computation, λ had even less influence on the remaining parameters. Interestingly, here μ was comparably strongly correlated with both A and AUC, and the correlation between A and AUC from the spline (0.854/0.963) was much higher than for the model. That is, λ was on average less strongly correlated with the other parameters in the case of the spline, whereas all other correlations were stronger. Accordingly, estimates for AUC correlated most strongly between model and spline, followed by μ, A and finally λ in

decreasing order. In figure 2 of the appendix to this chapter, graphical representations of the overall relative behaviour of the parameter estimates in all-against-all correlation plots for both the model and the spline fit approach are provided.

In Figure 3 eight examples for the distinct types of curve shapes identified in our datasets (wells G11 and H11 for all four strains, respectively, see Figure 1) are used to explore the specific behaviour of, and the potential problems specifically associated with, model fitting and spline fitting in comparison with the native OmniLog® PM software.

On substrate H11 (Figure 3, upper row) the respiration curves for all strains indicated positive reactions, but their shapes were substantially different. By using the curve parameters together with their CIs, the differences were easily detectable and one could intuitively comprehend the differences in curve shape. Spline fitting yielded broader CIs for λ and μ. None of the parameter estimates were biologically unreasonable.

On substrate G11 (Figure 3, lower row), besides two common curve shapes (blue and black) two deviating data situations are shown (red and green curve), which were nevertheless common in our measurements. The red curve reveals a primary, steep ascent followed by an interim plateau, before a second, shallower ascent conducts to the final maximum height. This behaviour occurred in all ten repetitions of this experiment. The model-based estimate for μ was lower than the spline-based one, but both were higher than the OmniLog® estimate.

The green curve describes an intrinsically negative reaction (no respiration curve), but instead a slight and linear increase in colour development. This probable noise was apparently sufficient as a data basis on which model fit was possible, but some of the resulting parameters, especially the negative length of the lag phase λ, were not biologically sensible. In contrast, the corresponding spline fit yielded a positive value for λ with a broad CI. Again in contrast, the OmniLog® software yielded a value near zero for this parameter. The estimate for μ was slightly positive only from the spline fit, whereas the other two methods yielded zero values. The other parameters were rather similarly estimated by the different methods.

To further explore the causes of the respective differences and characteristics of parameter estimation in certain data situations, the best-fitting model and the spline fit for the red and green curves from well G11 are shown in Figure 4. In case of the red curve (Figure 4, left side), the modelling method found a Gompertz exponential model to describe the data. However, the estimator for the maximum height A from the model approach is much too small, while both the OmniLog® and the spline estimates fit well to the real maximum height of the curve (see Figure 3, G11). Apparently the model fit is influenced too much by the

height of the non-typical interim plateau. In contrast, the spline was able to model the irregularity and hence to represent the curve's behaviour more precisely.

The example corresponding to (almost) no respiration (Figure 4, right side) was somewhat more complicated. Ideally, non-respiration would result in a horizontal line, and, hence, non-convergence for modelling approaches. However, the linearly increasing noise allowed a model to be fitted to the data which apparently resulted in biologically unreasonable parameter estimates *via* extrapolation; i.e., in the model the lag phase was extended to prior the beginning of the measurement (at 0 hour). In contrast, the spline approach was able to follow the data more precisely, apparently without the need to extrapolate.

Detecting significant differences

In Figure 5 the curves from ten technical repetitions of the reaction on substrate D12 (Minocycline) are compared with their curve parameters and 95% CIs estimated using the spline approach. These curves only differed regarding the beginning of the respiration reaction.

We used the red-coloured curve (D12/4) as an exemplar for demonstrating the detection of significant differences *via* CIs, which are indicated by vertical blue lines in the graphic. The two curves D12/1 and D12/2 were different to a statistically detectable degree regarding the length of the lag phase λ with a mean longer λ of 4.6 h and 12 h, respectively. D12/5 and D12/6 exhibited significantly larger slopes μ, differing in mean 15 and 14.5 units, respectively. Due to the very narrow CI for the maximal respiration A, D12/6 was identified as statistically detectable different with on average 3 OmniLog® units more respiration. D12/3, D12/7 and D12/9 had a smaller A with mean differences of 1, 2 and 1.7 units, respectively. Although all differences were statistically detectable, the user had the additional information of the effect sizes and thus was, in principle, able to use background information to decide whether the detected differences were biologically relevant. The integrals describing the areas under the curves resulted in very small CIs and thus all curves, except D12/3, were differing significantly.

The results from the time series approach in the third dataset are shown in Figure 6 for substrate C08 (L-Rhamnose). Curve 20, the fourth repetition from time point 21.92 h (t_5) was chosen as an example and the corresponding CI limits highlighted. For both λ and μ, all other CIs overlapped with that from curve 20, indicating no detectable differences between the curves. Considering the maximal respiration A and the integral AUC, several CIs did not

overlap with that from curve 20, but the effect size for the maximal respiration is at most 4 OmniLog® units (= 1.5%) for A and 978 units (= 5%) for AUC. Again, the user was now free to decide whether these differences were biologically relevant.

Regarding the comparison of group means, Figure 7 shows both the preliminary visualization using the mean CIs calculated over the groups (upper part) and the CIs for the differences between the means resulting from the simultaneous comparison procedure (lower part). The mean CIs can be used analogously to the strategy described above: overlapping CIs indicate no detectable difference between the groups, while non-overlapping ones indicate such differences. The testing procedure also provides 95% CIs for the differences between the group means, thus yielding precise information about the significance of the differences between the groups regarding the considered parameter(s).

To examine whether it is valid to subtract the negative controls (A01) from the measurements from all other wells before estimating curve parameters, we compared the parameter values for maximum height (A) from the A01 with that from selected wells with a negative reaction. Our findings suggests that the negative control might display a reproducible, strain-specific growth-like behaviour, and even though these curves are shallower than unambiguously positive reactions, their maximum height can can well be larger than that of typical negative reactions on the same plate. This makes it impossible to regard it as an approximation of an error term to be subtracted from the measurements from each other well. These findings are described in detail in the appendix to this chapter.

Which, and how many, classes of curve shapes?

Analysis of archetypes (Figure 8) indicated that either four or five archetypes are optimal. For five as predefined number, the resulting curve archetypes (insert in Figure 8) could be interpreted as follows: non-reaction with negative λ (an artefact, see above) (green line); non-reaction without such an artefact (black line); curves with a delayed start, i.e. reactions with a long lag phase λ, a relatively low μ and, thus, a rather low AUC/A ratio (blue line); early starting curves with a low λ, a moderate μ but nevertheless both high A and AUC (violet line); and, finally, rapidly accelerating curves with a moderate λ but a high μ, which reach an almost as high A and AUC (red line). These rapid accelerators had approximately the same A/AUC ratio as the early starters, but occurred much more seldom in the datasets (Figure 8).

Discussion

Visualization of PM raw data

When facing huge and complicatedly structured datasets such as the PM ones discussed here or that commonly occurring in other -omics analyses, the only way to get a comprehensive insight into the experimental results is a suitable graphical raw data representation. Such exploratory graphics have to be comprehensible in short time but also be highly informative (Chen et al. 2008). The convenience of an exploratory graphical representation depends mainly on its flexibility. Hence, the graphics should be easily adjustable to individual users' requirements in order to enable them to discover potentially all interesting and important features of the data. In contrast to the severely limited options for comparing different strains on the same substrate in a single pre-specified plot, as provided by the OmniLog® PM software (Biolog 2009), the user needs to compose the data unaffiliated.

In this study we explored an open-source solution for these specifications and showed that curve kinetics offer a more powerful data visualization than heat maps (Jacobsen et al. 2007). By using the function *xyplot()* from the *lattice* package (Sarkar 2008) highly structured graphics can easily be produced while retaining flexibility by systematically decoupling the various elements of a display. Itemization by substrate, tested strain or even repetition number was quite simple and constraints regarding the number of displayed curves or the position of the subpanel were not imposed at all. We thus recommend this or equivalent visualization approaches for PM data. A potential improvement compared to Figure 2 is the inclusion of the names of the substrates in addition to or instead of the mere coordinates of the wells.

Parameter estimation from respiration curves

The information content of the longitudinal PM raw data is a multiple of what an endpoint measurement could ever provide. A suitable analysis strategy thus has to be able to summarize this information and eliminate noise. These requirements can be met by model-fitting and spline-fitting approaches aiming on both dimension reduction and noise reduction (Fitzmaurice et al. 2004; Fitzmaurice and Ravichandran 2008). With *grofit*, the result is a set of four parameters sufficient for comprehensively describing the curves' shape. The main goals of a subsequent data evaluation would be the determination of the influence of different substrates, organisms investigated, or pretreatments, *via* the comparative characterization of respiration over time.

Although the OmniLog® PM software (Biolog 2009) is, in principle, able to provide a compilation of parameters, their computation is based only on few data points, potentially leading to the neglect of relevant data features. Here, we applied two alternative methods for extracting the four curve parameters λ, μ, A and AUC. The aim was to find a robust estimation method that was able to deal adequately with curves' potential deviations from the common sigmoid shape (Eilers and Marx 1996). It turned out that the spline smoother is flexible enough to follow even extreme curve shapes and is therefore superior for general parameter estimation, while the model-fitting approach appeared to be more constrained by the underlying model equations and straightened the curves to much. While 14% to 28% of the estimates for λ were biologically unreasonable in a strict sense (negative), most of these were only slightly negative and could safely be regarded as minor mis-estimates for 0.0. Also, the interpretability of the other parameters was not affected in these cases, and extremely negative λ can still be qualitatively interpreted as indicative of overall negative reactions.

The high amount of negative estimates for λ suggests that there is still space for improvement. In this study, the default parameters for the smoothing spline and the number of knots were used, since the evaluation of best-performing parameters was beyond the scope of this study. However, the selection of these two kinds of parameters is the critical step in this method (Eilers and Marx 1996). Also, other spline families and generalized additive model frameworks would exhibit interesting features for curve fitting by imposing monotonicity constraints on smooth effects and on ordinal, categorical variables (Hofner et al. 2011). We cannot exclude that as yet unimplemented models would outperform the ones considered here or even the spline fit, but in the current situation we regard the use of splines as the best recommendation that can be provided to users interested in fitting PM curves with R.

Compared to both model and spline methods, the slope estimates from the OmniLog® PM software (Biolog 2009) tend to be lower if the underlying curve is not perfectly sigmoid-shaped or the respiration reaction is finished long before the measurement is stopped (G11 in Figure 3, red and black curve, respectively). Since the OmniLog® PM software assumes that any reaction is symmetric around the inflection point, the slope is underestimated in the case of a secondary increase, which extends the distance between the time of inflection and the one of the maximum height. In contrast, the AUC estimates by the OmniLog® PM software (Biolog 2009) are slightly larger than those by the spline and model approaches, particularly for steep curves (Figure 3). As the native PM software represents the curves as series of rectangles, this deviation is most likely an overestimation and is expected to increase if more steep curves are encountered. Based on these results we favour the spline-based approach to

parameter estimation over the native PM software not only because it provides CIs but also because its point estimates are more robust against the described irregularities in curve shapes.

It is well known that phenomena such as autocorrelation (which is usual for growth curves) and non-homoscedasticity of the residuals violate the underlying assumptions of model- and spline-fitting (Härdle 1992; Eilers and Marx 1996). When dealing with high-throughput datasets such as the PM ones, however, the detailed checking of a potential violation of the assumptions made when fitting each curve is not practicable. Moreover, while for instance the spline might overfit the data in such situations, it is here used for smoothing each curve before extracting just four abstract parameters of interest. It is thus unlikely that potential violations of the underlying assumptions of the fit adversely affect the unbiasedness of the parameter estimates. This might explain why the spline appears more robust than the other methods if applied to PM data. While the assumptions of *ad hoc* approaches such as those implemented in (Biolog 2009) are, in general, less explicit, is is nevertheless apparent that they are frequently violated, too (Figure 3).

Detecting significant differences

In order to enable the user to extract all necessary information, we provided a feasible graphical solution displaying the point estimator together with its CI limits. The function *xyplot()* from the package *lattice* (Sarkar 2008) already provides the here presented outputs; only little adaptation of the input data is necessary (but see below). The straightforward assembly of different curves' characteristics in a single overview facilitates the interpretation and comparison of user-defined data subsets arranged according to technical and/or biological repetitions or other aspects of the experimental design.

With two exemplars (Figure 5 and 6) we familiarized the reader with the application of CIs to PM data for detecting (in-)significant differences. The demonstrated tool yielded valuable information about the range of variability of each point estimator on the corresponding scale. Thus, the user was enabled to recognize statistically detectable differences which he could further interpret regarding the specific biological relevance in each individual question. With the example in Figure 6 we demonstrated a further important approach. If, conversely, the experimenter wants to corroborate that a difference between the curves is not detectable (by defining a threshold *a priori* as the maximum allowed difference between the respective parameter estimates), the CIs provide a comprehensible solution (by allowing one to assess whether the expected mean difference is significantly larger than the threshold). This allows

one to assess whether reproducibility is given or whether the experimental procedure needs improvement, which is important for industrial applications or for research questions aiming at the identification of strains according to their metabolic features. For instance, Figure 6 shows that the dependency on the time of growth is negligible for this combination of organism and well and, hence, the protocol needs not be further standardized regarding the duration of growth.

It may often be of interest not to compare single curves but distinct groups of curves. In Figure 7 an example for the comparison of experimental group means, which is the method of choice in data evaluation for most biological questions, was shown. Starting with the preliminary method of mean CI calculation and their graphical representation, the user is encouraged to uncover interesting data features based on impartial calculations. Using the more sophisticated simultaneous calculation of differences of user-defined means in combination with the visualization of their CIs, the experimenter is empowered to investigate the data set more specifically regarding the biological hypotheses. The advantages of simultaneous CIs for drawing testing decisions are that the significance, relevance, and direction (increase or decrease) of the effect of interest, as well as the uncertainty concerning the estimates, can be interpreted in a scale close to that of the measured variable, which is often easier than interpreting p values in the scale of probability (Schaarschmidt and Vaas 2009).

Which, and how many shape classes of curve shapes?

The conducted archetype analysis (Eugster et al. 2009) indicated that only two classes of curve shapes are suboptimal, even if one corrects for the fact that at minimum two classes are necessary to represent the non-reactions alone due to the negative estimates for λ. Two to four archetypes were necessary for optimally representing the positive reactions, apparently because of fundamental differences in curve shape with a rather straightforward interpretation (Figure 8). Since the number of necessary archetypes depends on the analysed dataset, it is currently hard to recommend a predefined number of classes or even a rule of thumb. Larger datasets with even more distinct curve shapes might require more archetypes according to the RSS criterion, whereas biological background information might indicate even distinct numbers of categories. Anyway, the application of archetypes presented here already shows that a biologically meaningful post-processing of PM measurements *via* the parameter estimates is possible. Of course, other classification algorithms could also be applied such as k-means partitioning or even hierarchical clustering (Legendre and Legendre 1998). Even if

only the discrimination between positive and negative reactions was of interest, automatically classifying the observed curve shapes by assigning them to predefined clusters of curves or "typical" curves would be necessary for high-throughput processing of the Phenotype MicroArray data.

Conclusions and outlook

With the here presented approach to OmniLog® PM data analysis highly structured graphics can easily be produced while retaining the flexibility of systematically decoupling the various elements of a display. Itemization by substrate, tested strain or even repetition number is quite simple and no constraints about the number of displayed curves or the position of the subpanel are imposed at all.

The smoothing spline method for dimension and noise reduction appears to be more robust than model fitting when applied to PM data. *Via* curve-fitting the user can extract more information from the same experimental data than with any of the previously established techniques, particularly those reducing the data to binary states (positive vs. negative reactions). The inferred parameters can be used to classify the curves, and with our dataset more than just the categories positive/negative were optimal, even though the resulting archetypes could be easily interpreted. Dimension reduction of the curves followed by automated classification and identification seems to be of high future potential, particularly if combined with CIs, for the computational high-throughput processing of the raw data. This kind of data treatment will most probably also enhance the usefulness of high-throughput phenotyping for data modelling in microbial pathway genomics (Oh et al. 2007).

In conjunction with the proposed parameter estimation using models or preferably splines, the experimenter is free to define limits within which statistically detectable differences are not considered biologically relevant, which makes this method easily adaptable and more powerful than conventional mean comparison procedures (Schenker and Gentleman 2001; Everitt and Hothorn 2010). The proposed method intentionally resigns any multiplicity adjustment, because the analyses are intended to find preferably all interesting phenomena.

Although the strategy introduced here depends on CIs calculated on the basis of single curves, the approach could be easily extended to include the calculation of mean curves and corresponding intervals (Ritz and Streibig 2005), or to summarize the parameters from associated curves and perform CI computations and comparisons of multiple means with the resulting values (Everitt and Hothorn 2010). Considering the very low sample size in hitherto

published PM experiments, the chance to apply the hereby acquired information in sample size calculation should be emphasized. Usually one aim of statistical analyses is to find a detectable difference regarding an *a priori* chosen alpha (Everitt and Hothorn 2010). Since statistical testing provides primarily this detection of statistical significance, researchers frequently interpret only this information, irrespective of the size of detected mean differences, i.e. the effect size (Gardner and Altmann 1986; Cumming et al. 2007). However, for the majority of experimental investigations, especially in physiology (Baba et al. 2008; Edwards et al. 2009; Abu-Asab et al. 2011), often a minimum effect size is known for an effect to be biologically relevant. Our approach can extract information from preliminary experiments that can be used to compute the specific sample size required for the detection of biologically relevant differences with a sufficiently high confidence level in subsequent experiments. Thus, experimenters are enabled to improve their experimental design for satisfying their specific constraints and requirements more thoroughly.

Even though the analysis of the biological reasons for the respiration behaviours of the here tested strains is beyond the scope of the study, a few remarks on the study design and implementation of controls should be placed. On the GEN III plates the A01 well is defined as the control well, containing no substrate. By construction no reaction should occur on this well unless some kind of artefact was involved. The vendor's recommendation is, understandably, to adjust the experimental procedure until this point is met (Biolog 2009). However, further the user is instructed to subtract the, hopefully low, A01 curve from all other curves before proceeding with data analysis. Considering the fact that growth curves are seldom strictly additive or multiplicative in a biologically meaningful sense (Zimmermann and Nunez-Anton 2001), this approach raises several concerns regarding the impact on the shapes of the resulting curves and the character of thereby introduced biases. From our point of view, the experimental conditions should be first tried to be customized until there is no detectable positive reaction in A01 any more (Haack et al. 1995). We strongly encourage users to use the raw data for further analysis without subtraction of A01 from all other curves. We believe that the curve from A01 and its parameters are of higher benefit when used as thresholds for the dichotomization of experimental outcomes. The only exception would be a scenario in which the values in the negative control could be regarded as some kind of background noise which actually behaves additively with respect to the signal from the curves, if any. Our observations disagree with this scenario, however, as many intrinsically negative reactions in other wells were shallower (i.e., showed lower values of A) than those in A01. Surprisingly, the shapes of the curves were strain-specific, and for *Escherichia coli*

DSM 30083^T, if pooled over all replicates, the values of A in well A01 turned out to be significantly larger than those in, e.g., well D03 (see the appendix to this chapter).

In our assessments of the PM technique we observed a series of other experimental sources of errors. One of them is a false-positive colour development due to some chemical conversion of the redox dye, actually not caused by respiration (data not shown). Especially pentose sugars such as L-arabinose or xylose might be susceptible to these reactions (B. Bochner, pers. comm.). Thus, we recommend to measure one plate inoculated with only the inoculation fluid but no cell material and to check if such false-positive reactions occur. Wells with such reactions should be excluded from further analysis, since their colour development cannot safely be attributed to a physiological reaction.

To conclude, we believe to have demonstrated that tools provided in the free statistical software environment R can be successfully applied to PM data. These tools allow the user to visualize the kinetics in several meaningful respects, to conduct parameter estimation and, hence, dimension and noise reduction with the measurements and to detect statistically significant differences between the curves. All of these techniques can be conducted after selecting and rearranging the data in a sensible way depending on the respective scientific questions of interest. The outcome can even be used to improve the experimental setup itself as, e.g., by determining the necessary minimum number of replications. Additional work is necessary, however, to optimize details of the parameter estimation procedure (see above), and particularly to integrate all mentioned tools together with data input and output, and addition of metadata, in an easy-to-use R package. We are currently in the process of completing such a package, ‘opm’, and submitting it to the comprehensive R archive network CRAN (http://cran.r-project.org/).

Literature

Abu-Asab MS, Chaouchi M, Alesci S, Galli S, Laassri M et al. (2011) Biomarkers in the age of omics: time for a systems biology approach. OMICS: A Journal of Integrative Biology 15: 105-112

Akaike H (1973) Information theory as an extension of the maximum likelihood principle. In: Petrov BN, Csaksi F, editors. 2nd international symposium on information theory. Budapest: Akademia Kiado, pp. 267-281

Akaike H (1983) Information measures and model selection. Bulletin of the International Statistical Institute 22: 277-291

Baba T, Huan HC, Datsenko K, Wanner BL, Mori H (2008) The applications of systematic in-frame, single-gene knockout mutant collection of *Escherichia coli* K-12. Methods In Molecular Biology 416: 183-194

Becker RA, Cleveland WS, Shyu MJ (1996) A tour of Trellis graphics. AT & T Bell Laboratories, Murray Hill

BiOLOG Inc. (2009) Converter, File Management Software, Parametric Software, Phenotype MicroArray, User Guide, Part # 90333. Biolog Inc., Hayward

Birch C (1999) A new generalized logistic sigmoid growth equation compared with the Richards growth equation. Annals of Botany 83: 713-723

Bochner BR and Savageau MA (1977) Generalized indicator plate for genetic, metabolic, and taxonomic studies with microorganisms. Applied and Environmental Microbiology 33: 434-444

Bochner BR, Gadzinski P, Panomitros E (2001) Phenotype MicroArrays for high throughput phenotypic testing and assay of gene function. Genome Research 11: 1246-1255

Bochner BR (2009) Global phenotypic characterization of bacteria. FEMS Microbiological Reviews 33: 191-205

Bretz F, Hothorn T, Westfall P (2010) Multiple Comparisons Using R. CRC Press, Boca Raton

Brisbin IL, Collins CT, White GC and McCallum DA (1987) A new paradigm for the analysis and interpretation of growth data: the shape of things to come. The Auk 104: 552-553

Buchanan A V, Sholtis S, Richtsmeier J, and Weiss K M (2009) What are genes ”for” or where are traits “from”? What is the question? BioEssays 31:198-208

Bujara M, Panke S (2010) Engineering in complex systems. Current Opinion in Biotechnology 21: 586-591

Chang WE, Sarver K, Higgs BW, Read TD, Nolan NME et al. (2011) PheMaDB: A solution for storage, retrieval, and analysis of high throughput phenotype data. BMC Bioinformatics 12:109

Chen C, Härdle W, Unwin A (2008) Handbook of data visualization. Springer, Berlin

Christian BW and Lind OT (2006) Key issues concerning Biolog use for aerobic and anaerobic freshwater bacterial community-level physiological profiling. International Review of Hydrobiology 91: 257-268.

Cumming G, Fidler F, Vaux DL (2007) Error bars in experimental biology. Journal of Cell Biology 177: 7-11

Di Cagno R, Minervini G, Sgarbi E, Lazzi C, Bernini V et al. (2010) Short Communication: Comparison of phenotypic (Biolog System) and genotypic (random amplified polymorphic DNA-polymerase chain reaction, RAPD-PCR and amplified fragment length polymorphism AFLP) methods for typing *Lactobacillus plantarum* isolates from raw vegetables and fruits. International Journal of Food Microbiology 143: 246-253

Edwards RL, Dalebroux ZD, Swanson MS (2009) *Legionella pneumophila* couples fatty acid flux to microbial differentiation and virulence. Molecular Microbiology 71: 1190-1204

Efron B (1979) Bootstrap methods: Another look at the jackknife. Annals of Statistics 7:1-26

Eilers PHC and Marx BD (1996) Flexible Smoothing with B-splines and Penalties. Statistical Science 11: 89-121

Eugster MJA and Leisch F (2009) From Spider-man to Hero – archetypal analysis in R. Journal of Statistical Software 30: 1-23

Everitt BS and Hothorn T (2010) A handbook of statistical analysis using R. Chapman and Hall, Boca Raton

Feller W (1940) On the logistic law of growth and its empirical verification in biology. Acta Biotheoretica 5: 51-66

Feng X, Page L, Rubens J, Chircus L, Colletti P et al. (2011) Bridging the gap between fluxomics and industrial biotechnology. Journal of Biomedicine and Biotechnology 2010: 460717

Fitzmaurice GM, Laird NM, Ware JH (2004) Applied longitudinal analysis. Wiley, Hoboken NJ

Fitzmaurice GM, and Ravichandran C (2008) a primer in longitudinal data analysis. Circulation 118: 2005-2010

Fodor IK, Holtz-Morris AE, McCutchen-Maloney SL (2005) Growth curve models for the analysis of phenotype arrays for a systems biology overview of *Yersinia pestis*. Joint Statistical Meetings Minneapolis, MN, United States August 6, 2005 through August 10, 2005

Gardner MJ and Altman DG (1986) Confidence intervals rather than P values: estimation rather than hypothesis testing. British Medical Journal 292: 746-750

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biology 5: R80

Gottschalk PG, Dunn JR (2005) The five-parameter logistic: A characterization and comparison with the four-parameter logistic. Analytical Biochemistry 343: 54-65

Haack SK, Garchow H, Klug MJ, Forney LJ (1995) Analysis of factors affecting the accuracy, reproducibility, and interpretation of microbial community carbon source utilization patterns. Applied and Environmental Microbiology 61: 1458-1468

Hackett CA, Griffiths BS (1997) Statistical analysis of the time-course of Biolog substrate utilization. Journal of Microbiological Methods 30: 63-69

Härdle W (1992) Applied nonparametric regression. Cambridge University Press, Cambridge

Hofner B, Müller J, Hothorn T (2011) Monotonicity-constrained species distribution models ecology. Ecology 92:1895-1901

Homann OR, Cai H, Becker JM, Lindquist SL (2005) Harnessing natural diversity to probe metabolic pathways. PLoS Genetics 1: e80

Hütt MT (2001) Datenanalyse in der Biologie. Springer, Berlin

Jacobsen JS, Joyner DC, Borglin SE, Hazen TC, Arkin AP et al. (2007) Visualization of growth curve data from phenotype microarray experiments. 11th International Conference on Information Visualization (IV07), Zürich, Switzerland, July 4-6, 2007. Published by the IEEE Computer Society

Kahm M, Hasenbrink G, Lichtenberg-Frate H, Ludwig J, Kschischo M (2010). Grofit: fitting biological growth curves with R. Journal of Statistical Software 33: 1-21

Kandpal R, Saviola B, Felton J (2009) The era of 'omics unlimited. Biotechniques. 46: 351-355

Laird NM and Ware JH (1982) Random effects models for longitudinal data. Biometrics. 38: 963-974

Laird NM, Donnelly C and Ware JH (1992) Review Paper: Longitudinal studies with continuous response. Statistical Methods in Medical Research 1: 225-247

Lay JO Jr., Borgmann S, Liyanage R, Wilkins CL (2006) Problems with the "omics". Trends in Analytical Chemistry 25: 1046-1056

Le Cao KA, González I and Déjean S (2009) integrOmics: an R package to unravel relationships between two omics datasets. Bioinformatics Applications Note 25: 2855-2856

Legendre P, Legendre L (1998) Numerical ecology. 2nd ed. Elsevier, Amsterdam

Mayr E (1997) The objects of selection. Proceedings of the National Academy of Science USA 94: 2091-2094

Medina I, Carbonell J, Pulido L, Madeira SC, Goetz S et al. (2010) Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. Nucleic Acids Research 38: W210-W213

Mitchell DA, von Meien OF, Krieger N, Dalsenter FDH (2004) Review: A review of recent developments in modeling of microbial growth kinetics and intraparticle phenomena in solid-state fermentation. Biochemical Engineering Journal 17:15-26

Mithani A, Hein J, Preston GM (2011) Comparative analysis of metabolic networks provides insight into the evolution of plant pathogenic and non-pathogenic lifestyles in *Pseudomonas*. Molecular Biology and Evolution 28: 483-499

Neysens P, Messens W, Gevers D, Swings J and De Vuyst L (2003) Biphasic kinetics of growth and bacteriocin production with *Lactobacillus amylovorus* DCE 471 occur under stress conditions. Microbiology 149: 1073-1082

Nogales J, Palsson B and Thiele I (2008) A genome-scale metabolic reconstruction of *Pseudomonas putida* KT2440: iJN746 as a cell factory. BMC Systems Biology 2:79

Oh YK, Palsson BO, Park SM, Schilling CH, Mahadevan R (2007) Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data. Journal of Biological Chemistry 282: 28791-28799

Papin JA, Reed JL, Palsson BO (2004) Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. TRENDS in Biochemical Sciences 29: 641-647

Peugh JL, Enders CK (2005) Using the SPSS MIXED procedure to fit cross-sectional and longitudinal multilevel models. Educational and Psychological Measurement 65: 717-741

Puchałka J, Oberhardt MA, Godinho M, Bielecka A, Regenhardt D et al. (2008) Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology. PLoS Computational Biology 4: 1-18

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rabe-Hesketh S, Skrondal A (2008) Multilevel and longitudinal modeling using Stata. StataCorp LP, Texas

Reed JL, Patel TR, Chen KH, Joyce AR, Applebee MK et al. (2006) Systems approach to refining genome annotation. PNAS 103: 17480-17484

Richards FJ (1959) A Flexible Growth Function for Empirical Use. Journal of Experimental Botany 10: 290-301

Ritz C, Streibig JC (2005) Bioassay Analysis using R. Journal of Statistical Software 12: i05

Ruppin E, Papin JA, de Figueiredo LF and Schuster S (2010) Metabolic reconstruction, constraint-based analysis and game theory to probe genome-scale metabolic networks. Current Opinion in Biotechnology 21:502-510

Sabet S, Diallo L, Hays L, Jung W, Dillon JG (2009) Characterization of halophiles isolated from solar salterns in Baja Califonia, Mexico. Extremophiles 13:643-656

Sarkar D (2008) Lattice: multivariate data visualization with R. Springer, New York

Schaarschmidt F, Vaas L (2009) Analysis of trials with complex treatment structure using multiple contrast tests. HortScience 44: 188-195

Schenker N, Gentleman JF (2001) On judging the significance of differences by examining the overlap between confidence intervals. The American Statistician 55: 182-186

Serroyen J, Molenberghs G, Verbeke G, Davidian M (2009) Nonlinear models for longitudinal data. The American Statistician 63: 378-388

Sikorski J, Nevo E (2007) Patterns of thermal adaptation of *Bacillus simplex* to the microclimatically contrasting slopes of 'Evolution Canyons' I and II, Israel. Environmental Microbiology 9: 716-726

Singer JD (1998) Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics 23: 323-355

Spiegelman, D, Whissell, G, Greer, CW (2005) A survey of the methods for the characterization of microbial consortia and communities. Canadian Journal of Microbiology 51: 355-386

Sturino J, Zorych I, Mallick B, Pokusaeva K, Chang YY et al. (2010) Statistical methods for comparative phenomics using high-throughput phenotype microarrays. The International Journal of Biostatistics 6: 29

Ventura M, O'Flaherty S, Claesson MJ, Turroni F, Klaenhammer TR et al. (2009) Genome-scale analyses of health-promoting bacteria: probiogenomics. Nature Microbiol Rev 7: 61-71

Verbyla AP (1999) The analysis of designed experiments and longitudinal data by using smoothing splines. Applied Statistics 48: 269-311

Vilchez-Vargas R, Junca H, Pieper DH (2010) Minireview: Metabolic networks, microbial ecology and 'omics'technologies: towards understanding in situ biodegradation processes. Environmental Microbiology 12: 3089-3104

Warringer J, Blomberg A (2003) Automated screening in environmental arrays allows analysis of quantitative phenotypic profiles in *Saccharomyces cerevisiae*. Yeast 20: 53-67

Wenk MR (2010) Lipidomics: new tools and applications. Cell 143: 888-895

Wold S (1974) Spline functions in data analysis. Technometrics 16: 1-11

Xue X, Sztajer H, Buddruhs N, Petersen J, Rohde M et al. (2011) Lack of the delta subunit of RNA polymerase increases virulence related traits of *Streptococcus mutans*. PLoS ONE 6: e20075

Zhang W, Li F, Nie L (2010) Integrating multiple "omics" analysis for microbial biology: application and methodologies. Microbiology 156: 287-301

Zhou L, Lei X-H, Bochner BR, Wanner BL (2003) Phenotype micro-array analysis of *Escherichia coli* K-12 mutants with deletion of all two-component systems. Journal of Bacteriology 185: 4956-4972

Zimmermann DL, Nunez-Anton V (2001) Parametric modelling of growth curve data: An overview. Test 10: 1-73

Zwietering MH, Jongenburger I, Rombouts FM, van T'Riet K (1990) Modeling of the bacterial growth curve. Applied and Environmental Microbiology 56: 1875-1881

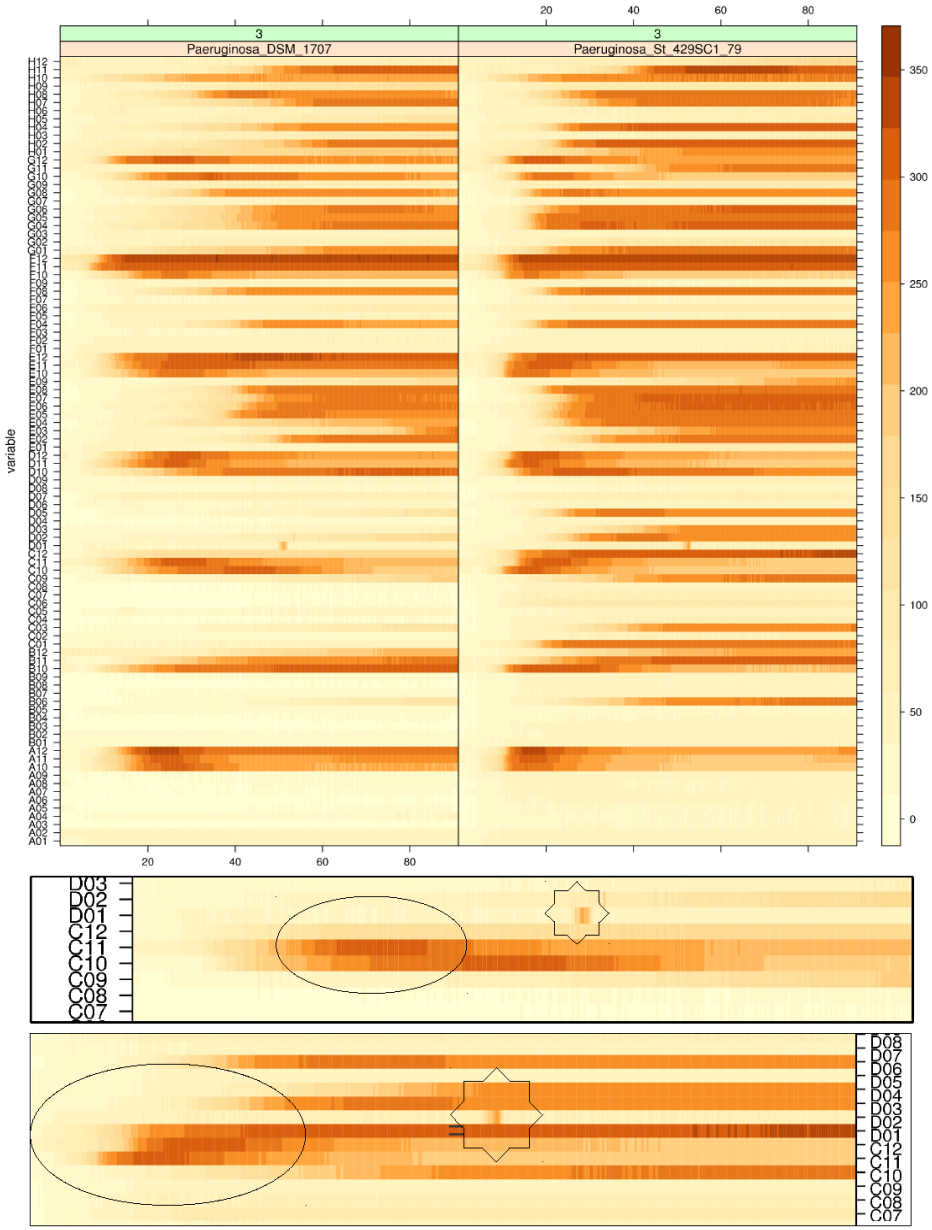
Tables

	Value	**Dataset 1**	**Dataset 2**	**Dataset 3**
curves	# total	3840	3840	3265
	# without fittable models	244 (6.35%)	64 (1.66%)	44 (1.35%)
	# without fittable splines	0 (0%)	0 (0%)	0 (0%)
experimental groups	# total	768	768	864
	# without fittable models	12 (1.56%)	1 (0.13%)	2 (0.23%)
	# without fittable splines	0 (0%)	0 (0%)	0 (0%)
	# model parameter $\lambda < 0$	281 (36.59%)	229 (29.82%)	205 (23.73%)
	# spline parameter $\lambda < 0$	221 (28.78%)	154 (20.05%)	124 (14.35%)
	# model parameter $A > 400$	10 (1.3%)	3 (0.39%)	12 (1.39%)
	# spline parameter $A > 400$	0 (0%)	0 (0%)	0 (0%)
	# model parameter $A < 100$	282 (36.72%)	245 (31.9%)	162 (18.75%)
	# spline parameter $A < 100$	291 (37.89%)	236 (30.73%)	146 (16.9%)
spline fits	# $A < 100$	1464	1196	583
	# $A < 100$ and $\lambda < 0$	884 (60.38%)	733 (61.29%)	485 (83.19%)
	mean λ if $A < 100$ and $\lambda < 0$	-207.6 ±2871.0	-6.8±8.1	-12.1±12.4
	# $A > 100$	2376	2664	2681
	# $A > 100$ and $\lambda < 0$	106 (4.46%)	106 (3.98%)	130 (4.85%)
	mean λ if $A > 100$ and $\lambda < 0$	-3.3±4.4	-3.3±4.3	-3.0±3.5

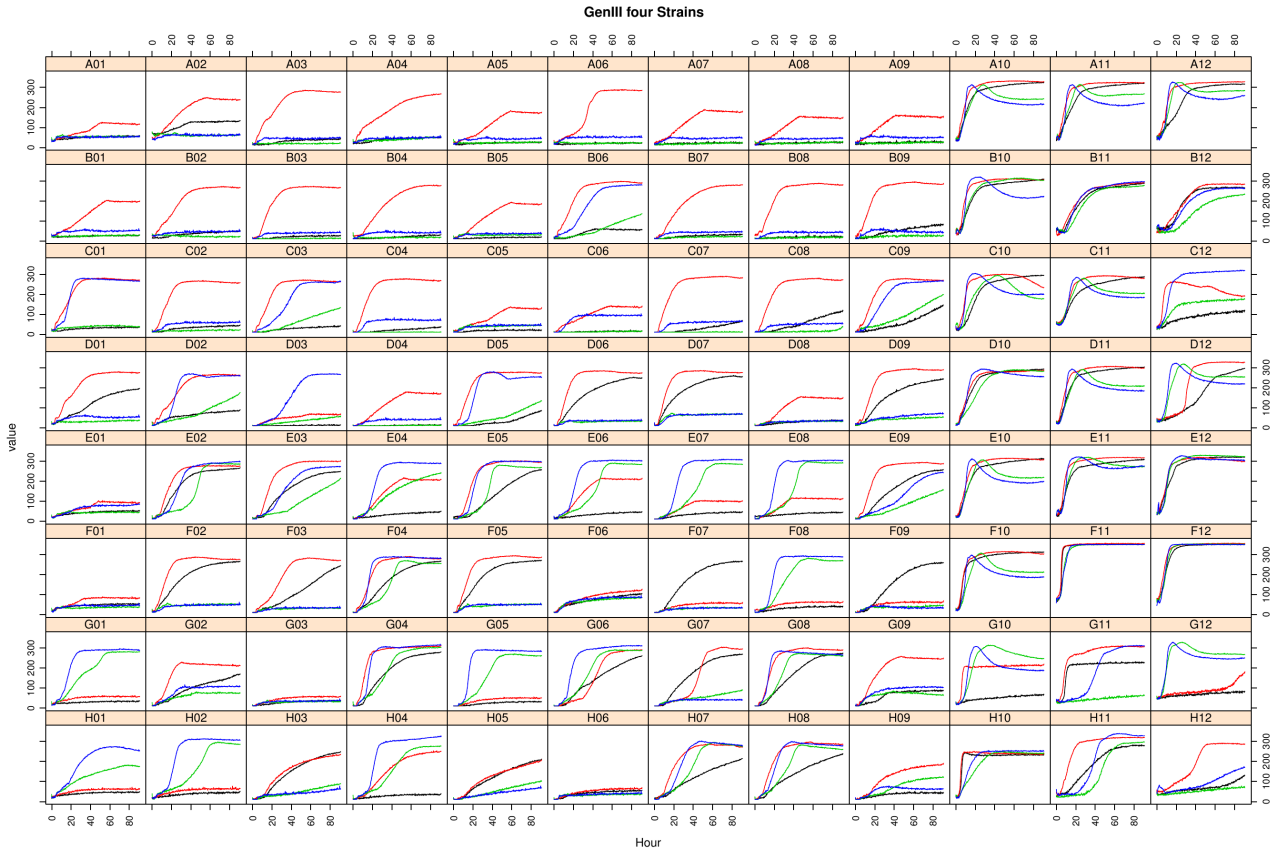
Chapter V, table 1. Robustness in parameter estimation. Summary statistics from parameter estimation from the three dataset exemplars using both the model fitting and model-free spline fitting approach from the basic part of R's add-on package *grofit* (Kahm et al. 2010). Results with parameters $\lambda < 0$ and $A > 400$ indicate biologically unreasonable estimates; parameter estimates $0 < A < 100$ approximately indicate negative reactions.

		Model				Spline			
		λ	μ	A	AUC	λ	μ	A	AUC
Model	λ	-	0.428	0.367	0.294	0.641	0.360	0.371	0.295
	μ	0.248	-	0.536	0.732	0.296	0.846	0.717	0.734
	A	0.362	0.320	-	0.7	0.324	0.538	0.748	0.7
	AUC	0.372	0.712	0.522	-	0.205	0.721	0.843	0.998
Spline	λ	0.571	0.131	0.225	0.221	-	0.299	0.285	0.203
	μ	0.940	0.837	0.272	0.620	0.025	-	0.723	0.736
	A	0.437	0.658	0.571	0.963	0.051	0.584	-	0.854
	AUC	0.372	0.713	0.523	0.999	0.046	0.621	0.963	-

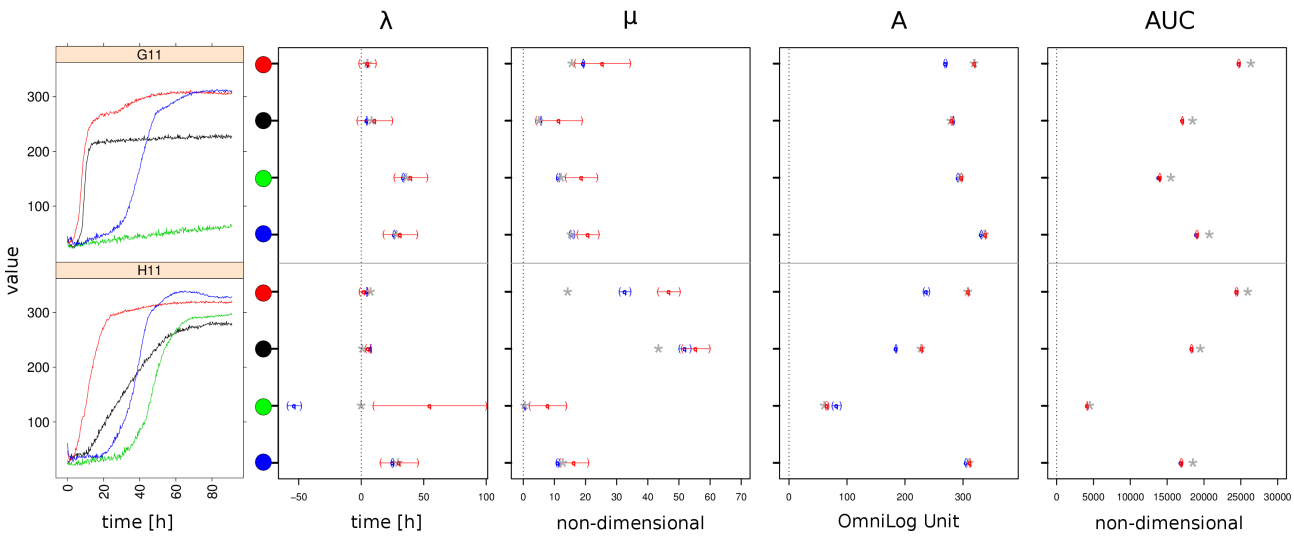
Chapter V, table 2. Within-method and between-method interdependence of parameter estimates. All-against-all correlations measured using Kendall's τ (above the diagonal) and Spearman's correlation index (below the diagonal) describe how the different curve parameters estimated using either model fitting or spline fit are associated with each other and with the corresponding parameters from the alternative fitting approach.



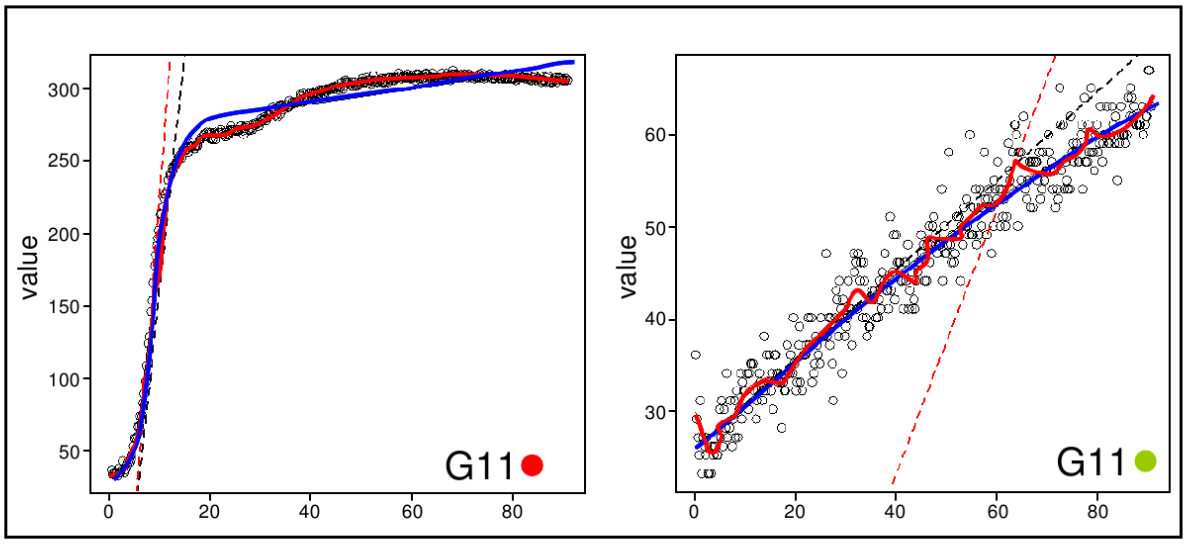
Chapter V, figure 1. Visualization of PM curves as heat maps *via* the function *levelplot()* as a re-implementation of the approach of Jacobsen et al. (2007) in R. Each respiration curve is displayed as a thin horizontal line, in which the curve height as measured in OmniLog® units is represented by color intensity (darker parts indicate higher values). The x-axes correspond to the measurement time in hours. The upper part shows an overview of two plates. Here, the descriptions of the y axes (only visible if enlarged, but see below) list the names of the wells; the descriptions of the x axes list the measurement time in hours. The boxes below represent magnified parts of the upper panel to illustrate the color changes in the case of decreasing color intensities (regions surrounded by black ellipses) or technical problems such as peaks (positions marked by eight-pointed stars).



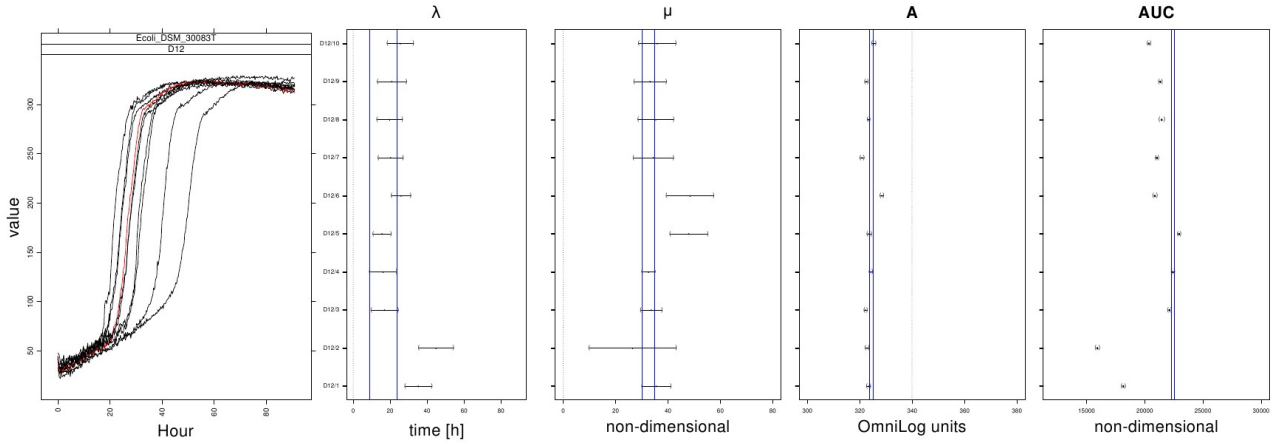
Chapter V, figure 2. Visualization of PM curves as such *via* the function *xyplot()*. PM curves from a representative technical repetition from the first dataset were arranged according to the original 8×12 wells plate layout. The respective curves from all four strains are superimposed; the affiliation to each strain is indicated by color as follows: black, *E. coli* DSM 18039; red, *E. coli* DSM 30038^{T}; green, *P. aeruginosa* DSM 1707; blue, *P. aeruginosa* 429SC. The x-axes show the measurement times in hours, the y-axes the curve heights in OmniLog® units. In the caption of each panel the corresponding coordinate of the well is shown. Details of the curves from wells G11 and H11 are examined in Figs. 3 and 4.



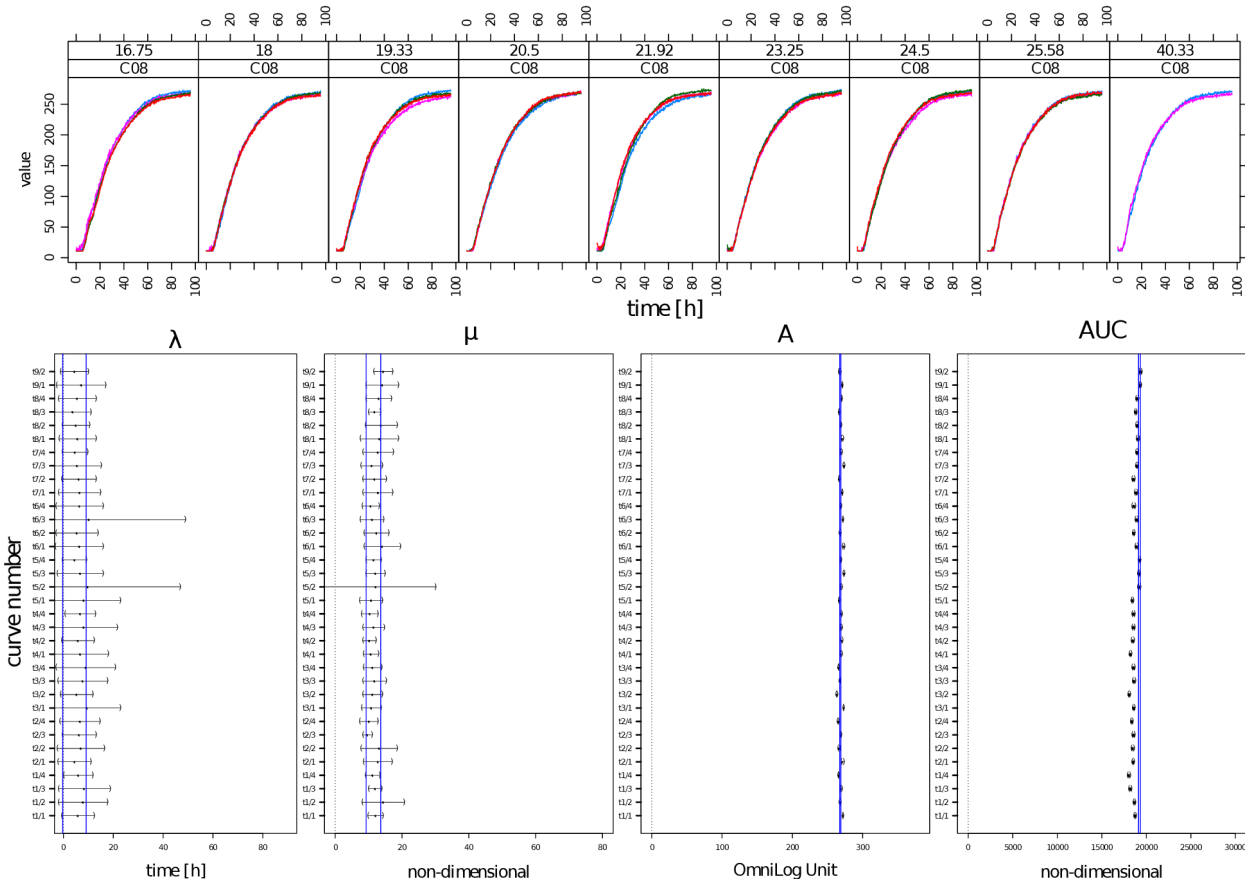
Chapter V, figure 3. Comparison of parameter and CI estimates from the same curves using three distinct approaches. Left, enlarged view of the curves from wells G11 and H11 as depicted in Fig. 2. As in Fig. 2, the affiliation to each strain is indicated by color as follows: black, *E. coli* DSM 18039; red, *E. coli* DSM 30038^{T}; green, *P. aeruginosa* DSM 1707; blue, *P. aeruginosa* 429SC. Right, point estimates and 95% CIs for each of the four parameters lag phase (λ), slope (μ), maximum (A) and area under the curve (AUC) estimated from the eight curves depicted on the left using either the model-fitting (blue dots and CIs) or the spline approach (red dots and CIs). The gray stars are the respective point estimators inferred with the native OmniLog® software (which does not provide CIs). The colored circular areas refer to the colors of the curves in the left part of the figure.



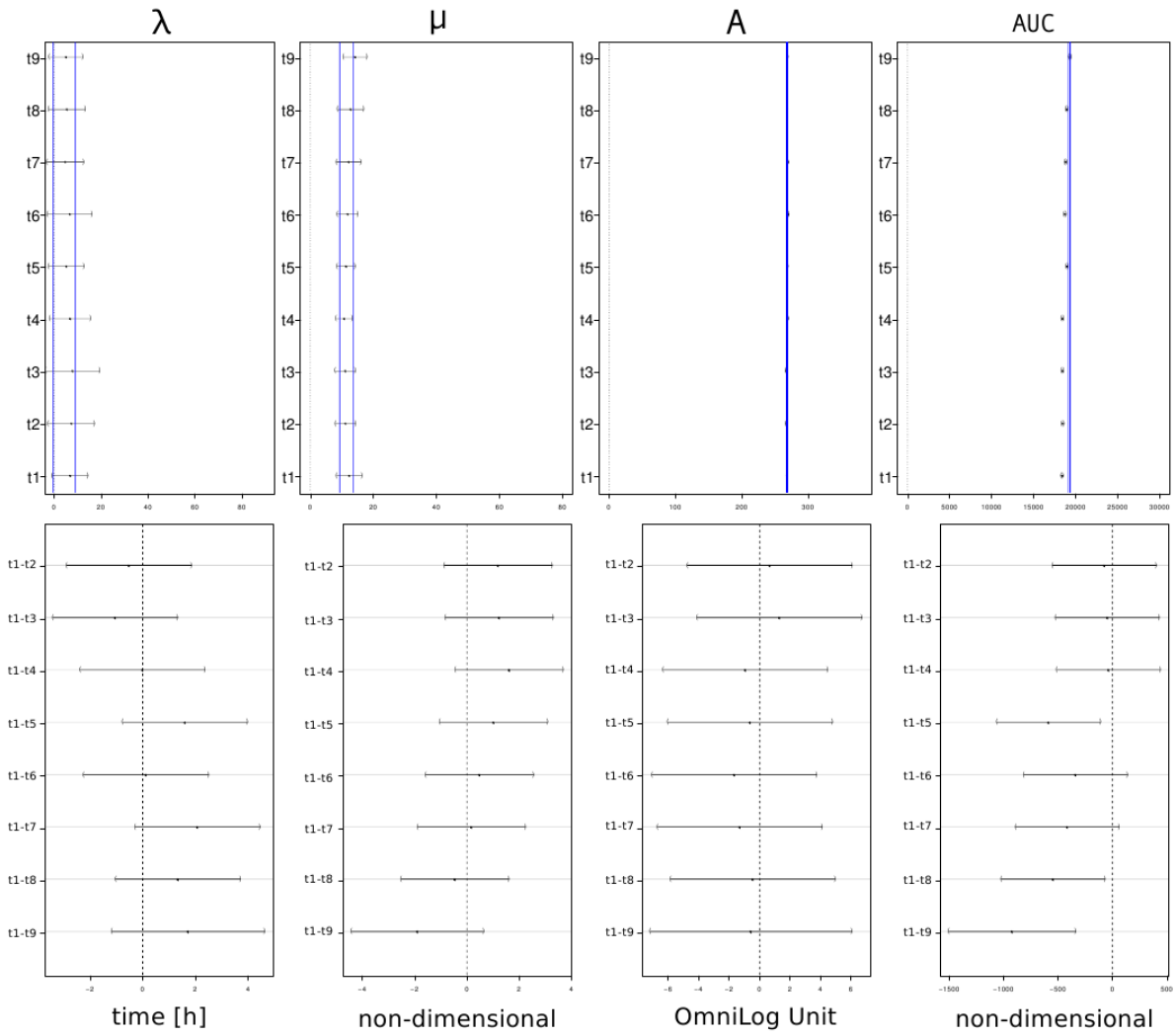
Chapter V, figure 4. Visualization of parametric and model free spline fit for the two special cases from G11, the red (here left) and green (here right) strain. The raw colour intensities (black circles), measured over time (x-axis, hours) were fitted by both a parametric model (thick blue line) and a model-free spline (thin red line). The thin dashed lines indicate the maximum slope of each approach (thin dashed black line corresponds to the model fitting approach, the thin dashed red line to the spline, respectively). In the left panel the irregularity is better customized by the spline fit, whereas the model straightens it with the consequence of overestimating the maximum height (A). In case of (almost) no respiration (right panel), the fitted model extrapolated the data and thus yielded biologically unreasonable parameter estimates. In contrast, the spline approach was able to follow the data more precisely, apparently without the need to extrapolate.



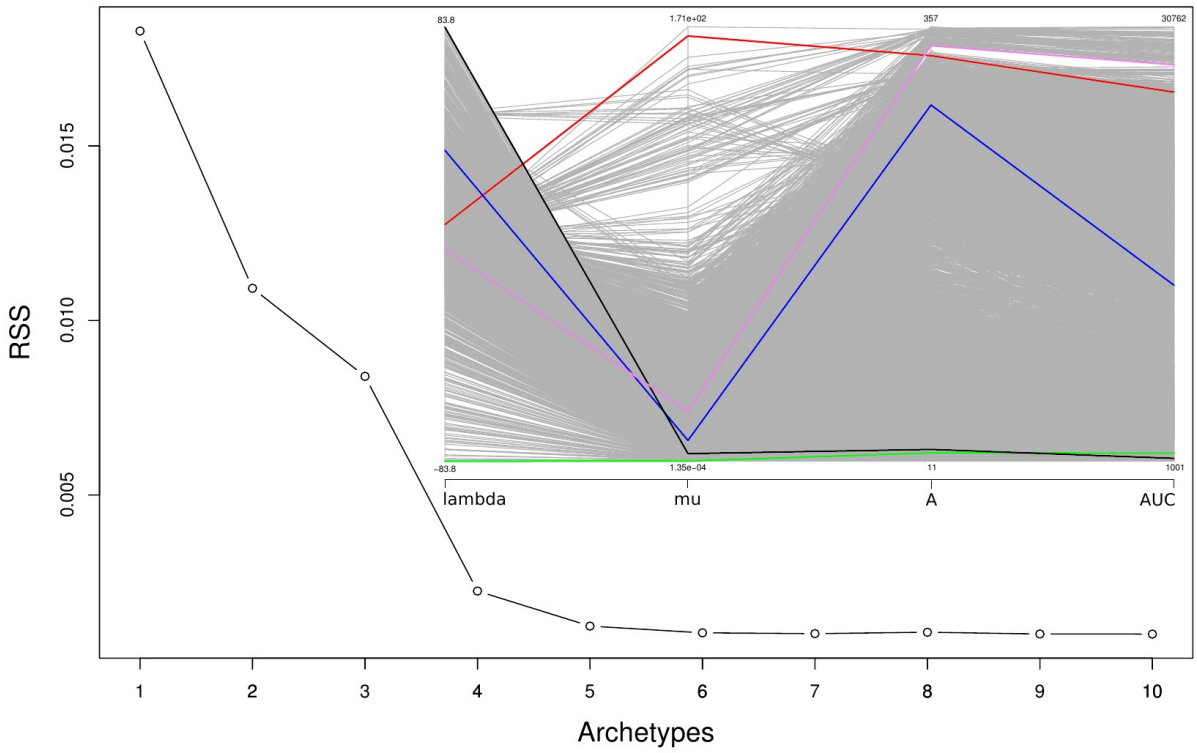
Chapter V, figure 5. Comparison of curve-parameter point estimates and their 95% CIs for each of the four parameters lag phase (λ), slope (μ), maximum (A) and area under the curve (AUC), estimated for ten technical repetitions of respiration on well D12. Left, a plot of the raw respiration data illustrates their courses individually for each of the ten repetitions. The red curve (D12/4) was used as an exemplar for demonstrating the detection of significant differences via CIs. In the right panel, point estimates and 95% CIs for each of the four parameters from the spline approach are given for each replication. The blue lines highlight the position of the upper and lower limit of CIs from D12/4's parameters. A non-overlap of the CIs of different curves indicates a difference of a statistically detectable amount, and the distance between two intervals provides information about the expected minimum difference.



Chapter V. figure 6. Comparison of point estimates and their 95% CIs for each of the four parameters lag phase (λ), slope (μ), maximum (A) and area under the curve (AUC), estimated for four technical replicates of respiration on well C08, in which the cells were additionally subjected to distinct pretreatments (cultivation times). The upper panels show the plot of the respiration curves of *E. coli* DSM 18039 = K12 on well C08 when grown on solid LB medium for 16.75 h (t1), 18 h (t2), 19.33 (t3), 20.5 (t4), 21.92 (t5), 23.25 h (t6), 24.5 h (t7), 25.58 h (t8) or 40.33 h (t9), respectively, and subsequently measured on GEN III microplates™ in the PM modus over 91 h. The mid plot shows point estimates and 95% CIs for each of the four parameters from the spline approach. The blue lines highlight the position of the upper and lower limits of the CIs from repetition no. 4 at t5. A non-overlap of the CIs of different curves indicates a difference of a statistically detectable amount, and the distance between two intervals provides information about the expected minimum difference.



Chapter V, figure 7. Visualization of group-wise representations of the four curve parameters lag phase (λ), slope (μ), maximum (A) and area under the curve (AUC). The upper panels show the results from the preliminary calculation, a simple calculation of group means of confidence limits and point estimators. The groups, here the distinct pretreatments (cultivation times t1 to t9), are given on the y-axis. For orientation, the blue lines highlight the position of upper and lower limit of CIs from repetition no. 4 at t5, in analogy to Fig. 6. In the lower panels the 95% CIs for the differences of group means are represented. The set of user-defined comparisons was calculated for the point estimators of each of the four parameters lag phase (λ), slope (μ), maximum (A) and area under the curve (AUC). Since these are CIs for the differences between the means, a non-overlap with zero indicates a statistically detectable difference between the considered group means of the examined curve parameters.



Chapter V, figure 8. Results from an archetype analysis of the four parameters estimated from the PM curves obtained from the 1st and the 2nd dataset using the smoothing splines. The outer figure is a scree plot in which the residual sums of squares (RSS, y-axis) are plotted against the corresponding predefined numbers of archetypes (x-axis). Apparently either four or five archetypes are optimal according to the "elbow criterion". The insert (upper right) is a parallel coordinates plot showing the original measurements (gray lines) as well as the optimal archetypes (green, black, blue, violet and red lines) obtained if five archetypes are requested. On the x-axis, the names of the curve parameters are indicated. The minima and maxima of the four y-axes are also indicated. For an interpretation of the archetypes, see the main text.

Chapter V, appendix

Behaviour of the negative controls compared to negative reactions in other wells

As mentioned in the main text of the chapter, it is frequently recommended to subtract the measurements from the well A01 (a negative-control well without any substrate) from those of the measurements from each other well *before* inferring curve parameters. We here (1) briefly recapture what we regard as the main underlying assumption of this recommendation and (2) discuss it in the light of some empirical observations with the datasets analysed in the main manuscript. The fact that subtraction cannot that easily be applied to the inferred parameters themselves has been highlighted there.

Background assumptions when subtracting the negative control prior to parameter estimation

The recommended procedure is to subtract the, hopefully low, A01 curve from all other curves before pursuing with data analysis. This strategy assumes a biologically sensible additivity between the negative control and respiration reactions caused by the substrates. This additivity can be formally expressed as follows. Let $v_{ij} \geq 0$ be the measured value in the i^{th} well at the j^{th} time point, $s_{ij} \geq 0$ be the hypothetical value cleaned from background noise at these coordinates, $e_{ij} \geq 0$ the hypothetical background noise in this position, and n be the position of the negative-control well. Apparently we have

$$s_{ij} = v_{ij} - e_{ij}.$$

The suggestion to subtract the values in the negative-control well implies that

$$e_{ij} \approx v_{nj},$$

and, hence,

$$s_{ij} \approx v_{ij} - v_{nj}.$$

Due to the non-negativity constraint of all s_{ij} this implies that

$$v_{ij} \geq v_{nj}.$$

Empirical behaviour of the negative controls in our datasets

Figure 1 compares the shapes of the negative-control curves between the four tested strains. Apparently the behaviour of well A01 is strain-specific. *E. coli* DSM 18039 and the two *Pseudomonas* strains display a typical negative reaction, whereas the type strain of *E. coli*

shows a more growth-like curve-shape (even though the maximum height is still low compared to the unambiguously positive reactions on the same plate).

Figure 2 confirms for *E. coli* DSM 30083^{T} that this typical shape of the negative-control curve occurs also throughout the 2nd biological replication. Moreover, we selected a well with a typical negative reaction, D03 (D-Arabitol), and compared it to the negative control. All D03 curves appeared shallower than the negative control.

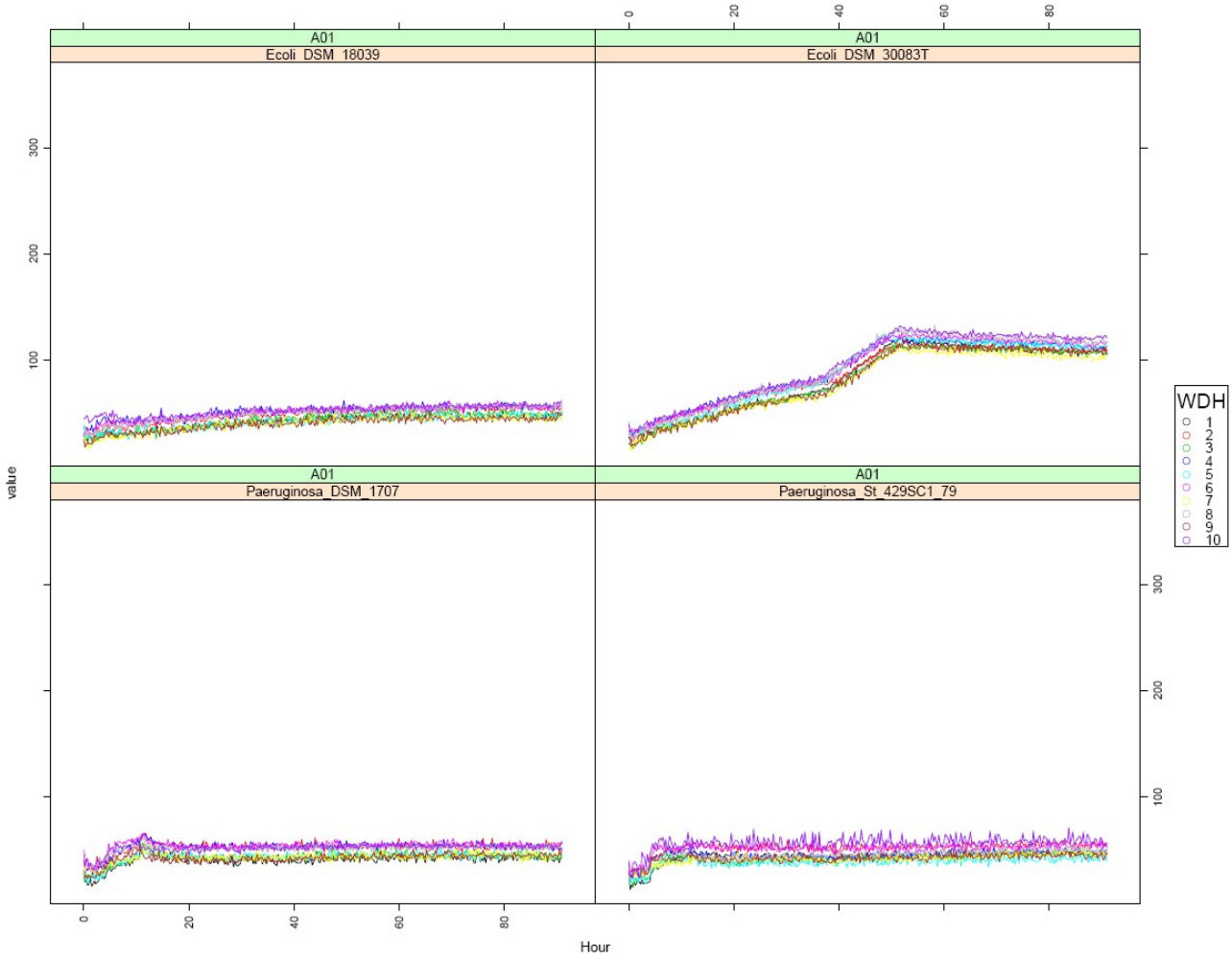
In order to statistically confirm this observation, we compared the parameter values for maximum height (A) from the negative control well A01 with that from well D03. For each dataset, a single one-sided t-test (test on decrease) with a confidence level of 95% was calculated, resulting in two statistically detectable group mean differences, p = 6,241e^{-13} for dataset 1 and p = 5,622^{e-10} for dataset 2. That is, D03 is significantly shallower than A01.

These empirical results are in sharp contrast to the above outlined theoretical assumptions which need to be fulfilled for subtracting the A01 values from the values of the other wells. They are therefore also in sharp contrast to the assumption of a biologically sensible additivity between the negative control and respiration reactions caused by the substrates. Rather, the negative control might display a reproducible, strain-specific growth-like behaviour. This makes it impossible to regard it as an approximation of an error term to be subtracted from the measurements from each other well.

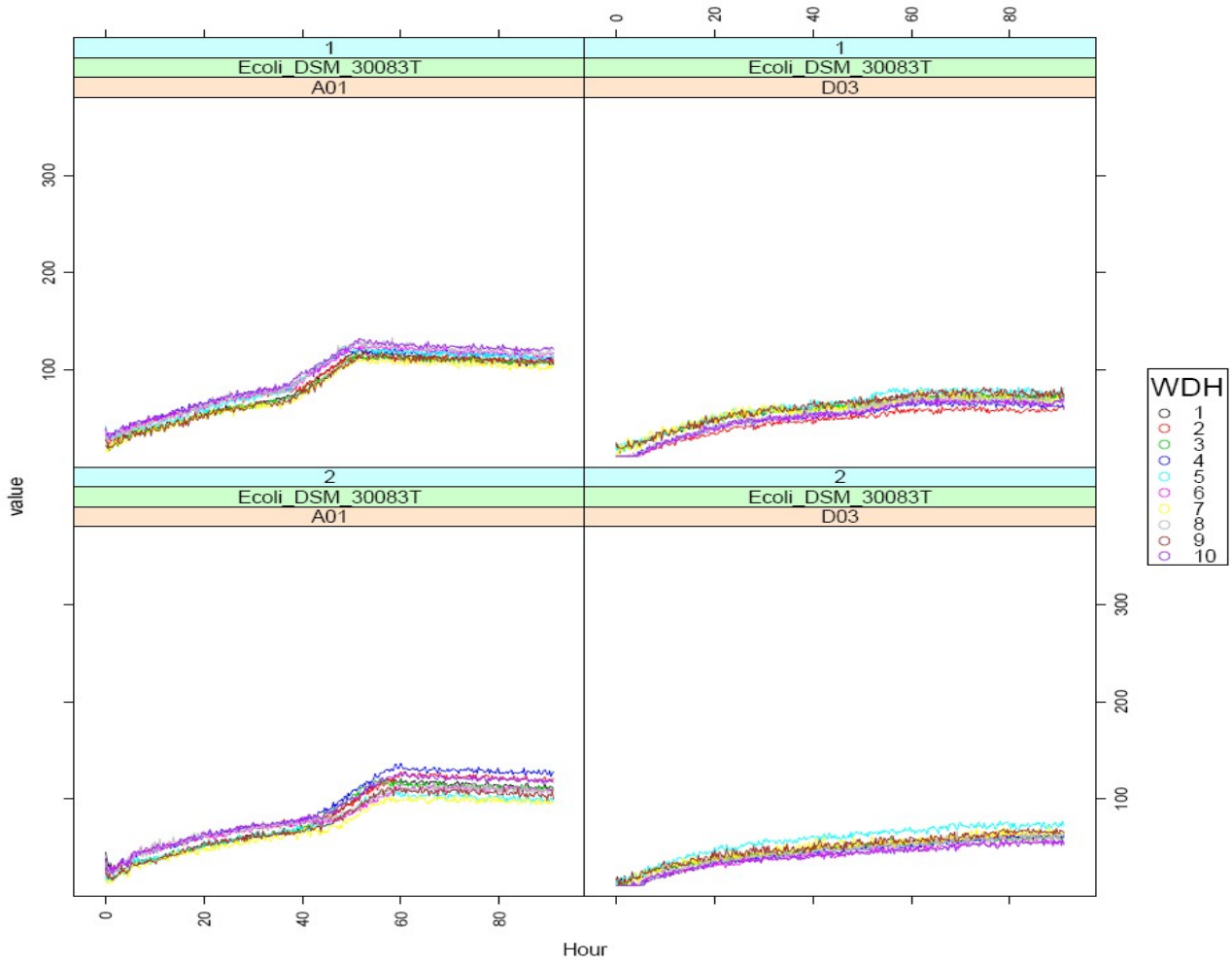
In such a situation, the choice a suitable strategy depends on the interpretation of the growth-like behaviour in well A01. One could either discard these results as due to a not yet sufficiently optimized treatment of the strain under consideration and try to modify the pretreatment and/or the composition of the incubation medium until the curves in A01 become shallower. Alternatively, the plates could be used a such, but when dichotomizing them into positive and negative reactions, a threshold would need to be chosen that yielded negative reactions in A01. All in all, our observations (see main manuscript) strongly argue for this 2nd alternative.

All-against-all correlation plots of the parameter estimates

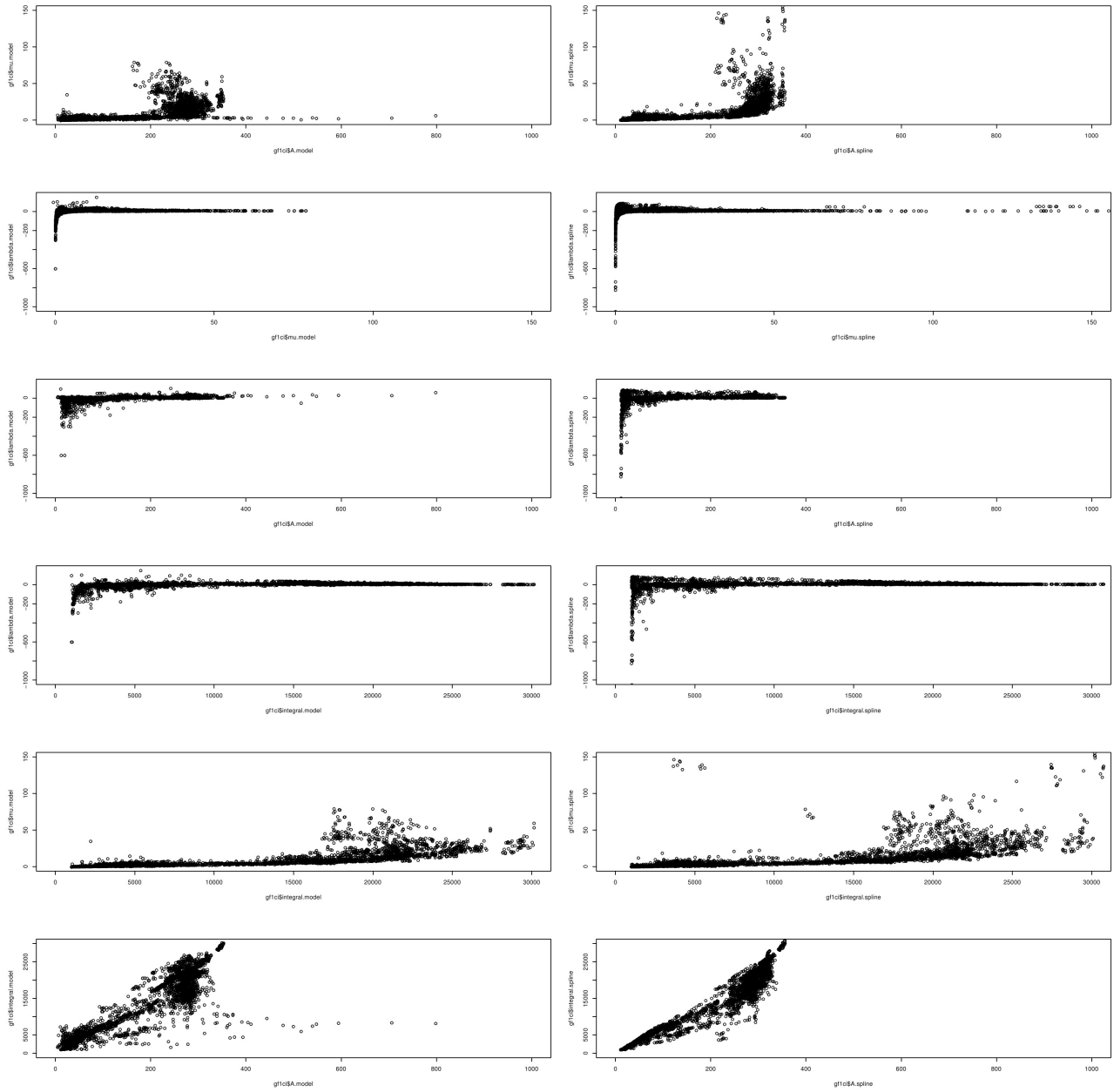
In addition to the all-against-all correlations of the parameter estimates listed in the main text of this chapter, corresponding plots are shown in Figure 3 of this appendix.



Appendix to chapter V, figure 1. Visualization of PM curves as such *via* the function *xyplot()*. PM curves from all ten technical repetition (WDH) from the first biological replication for the four tested strains on the well A01 (negative control) arranged according a 2×4 panel layout. In the caption of each panel the corresponding the strain name is shown. The x-axes show the measurement times in hours, the y-axes the curve heights in OmniLog® units.



Appendix to chapter V, figure 2. Visualization of PM curves as such *via* the function *xyplot()*. PM curves from all ten technical repetition from both datasets for the type strain of *E. coli* on the wells A01 (negative control) and D03 (D-Arabitol) were arranged according to a 2×4 panel layout. In the caption of each panel the corresponding dataset (biological replication 1 or 2), the coordinate of the well (A01 or D03) and the strain name is shown. The x-axes show the measurement times in hours, the y-axes the curve heights in OmniLog® units.



Appendix to chapter V, figure 3. All-against-all correlation plots of the parameter estimates.

VI. Discussion

The objective of this thesis was to explore how the homologous overexpression of *pr-10a* affects plant cells' stress responses. As biological system heterotrophic, dedifferentiated cell lines from *Solanum tuberosum* cv. Désirée (wild type and homologously *pr-10a* overexpressing ones) were chosen. The unusual framework of cryopreservation as well as classical osmotic and salt challenges constituted the settings for stress-response measurements. To cover the cells' reaction most comprehensively, monitoring was performed by molecular methods (transcriptomics) and on superordinate phenotypic levels (classical growth challenges and newly developed respiration kinetics).

Beyond the basic insights into the role of *pr-10a* in terms of plant cell stress response, three parts of the thesis report methodological improvements that outreach previous applications in the biological system under study. First, I showed that the specific combination of 1st cistron element and IRES element determines the success of the 2nd cistron target gene expression. Thus, I was able to formulate recommendations for devising testing steps and possible solutions for cases were the first cistron gene negatively effects the translational efficiency of the IRES element. Second, based on an experimental set-up originating from microbiology, I facilitated the sophisticated analysis of respiration kinetics via advanced visualization tools and statistical analysis software. The thesis clearly demonstrated that respiration measurements are a meaningful complement for the set of methods when examining phenotypes of heterotrophic dedifferentiated cell cultures. The novel method provides valuable information in situations not targeted by more common methods such as growth measurements. Third, I explored the free software environment R and devised a suitable analysis strategy for data measured over time, i.e. a comprehensive and adaptable framework for respiration kinetics data visualization and parameter estimation.

Below, I briefly summarize and discuss all parts of the thesis.

Stress response on the cellular level

Differential gene-expression analysis

The elaboration of the gene-expression patterns of *pr-10a* revealed detailed information about the complex interaction between the type of the applied stress (i.e., osmotic or salt challenge), its severity regarding osmotic pressure and its duration on the one hand, and *pr-10a* expression on the other hand. Although studies examining gene-expression patterns with a

comparably high resolution are, in principle, in demand when cellular stress response and signal transduction should be investigated, such sophisticated experimental designs are absolutely not common. Only the application of those intricate sampling strategies facilitates the disclosure of such well defined and timed expression patterns as demonstrated here for the wild-type cells, and their corroboration in the transgenic lines.

The experiments in this thesis were conducted with two independent transgenic cell lines. From the previously scheduled number of at least five independent cell lines homologously overexpressing *pr-10a* (El-Banna et al. 2010) those which were cryopreserved and stored in liquid nitrogen exhibited bacterial infections after thawing and during regrowth. Since the contaminants were not manageable by common rescue strategies using antibiotics, these lines were not available in time for the investigations in the course of this thesis. Further investigations might indeed employ additional independent cell lines for ensuring that the observed effects are not caused by genetic modifications during the transgenic insertion and following chromosomal reorganization processes. However, the two cell lines examined here exhibited a general consistency and similarity between both lineages.

When applying the cryopreservation framework, a threshold for *pr-10a* expression induction that determined cryopreservation success or failure was not revealed, but rather a correlation of transcript's incidence time and its continuance with cell survival. Together with the findings from the classical osmotic and salt-challenge experiments, it is supposed that a close correlation between the simple amount of transcripts and the effectual stress response does not exist. In fact, this fits to the previously postulated complex protein-regulation mechanisms, but at least to the role of *pr-10a* in signal transduction (Liu and Ekramoddoullah 2006) or in translating signals into the initiation of further stress-response reactions in the cells. Gathering information about incidence and magnitude of PR-10a protein abundances in stress challenges would be the next step for elucidating the role of PR-10a proteins in plant cells' stress-response cascades.

In the osmotic- and salt-challenge experiments also the transcription of the regulators *sebf* and *pti4* was monitored. Forming a heterodimeric SEBF-Pti4 complex, the two proteins were previously shown to disable the expression of *pr-10a* via the occupation of the silencer element of the promoter (Gonzales-Lamothe et al. 2008). For the activation of *pr-10a* gene expression, the complex has to dissociate from the DNA again, and the proteins are hypothesized to also dissociate from each other (Gonzales-Lamothe et al. 2008). Since in the wild-type cells the here applied osmotic or salt challenges caused an appreciably altered gene

expression for *pr-10a* but neither for *sebf* nor for *pti4*, it was shown that the applied treatments do not affect the transcription rates of these regulators. A possible explanation for the observed highly regulated *pr-10a* expression in the wild type cells is that after dissociation from the *pr-10a* promoter, the regulator proteins persist in the cells. Supposing that they are not being catabolized during the few hours of increased *pr-10a* expression, the proteins are still available to reconstitute the repressosome and silence the *pr-10a* gene again. Such a mechanism was not unexpected (Gonzales-Lamothe et al. 2008), and further research could focus on the details of the mechanism as, e. g., on protein turn-over.

Regarding the observed expression induction of *sebf* and *pti4* in both transgenic lines, together with the expression-induction patterns of the *luc* transcripts an interesting manifest coincidence was revealed. That is, similar to the salinity-dependent induction of the transgene-controlling mannopine-synthase promoter (p-MAS), the regulators *sebf* and *pti4* exhibited a well-structured expression pattern. Since the wild-type cells appreciably altered neither the *sebf* nor *pti4* gene expression under the sorbitol and the equiosmolal 0.32M NaCl treatments, but both transgenic lines did, it is supposed that the *sebf* and *pti4* expression induction in the transgenic lines is caused by altered transgene expression during stress. Actually, this would not be a severe problem in breeding frameworks focusing on improved phenotypes only, since this was accomplished anyway. With respect to our understanding of the underlying mechanisms, those events may provide interesting insights but raise questions as well. It follows from the findings of Gu et al. (2002) that *pti4* activates the expression of a wide array of pathogenesis-related genes and plays an important and distinct role in plant-defense mechanisms. Accordingly, it might be that the transgenic *pr-10a* disturbed the adjustment of plant-defense gene regulation under severe salt stress via the modification of the *pti4* and *sebf* expression levels and thus modified the effect of the its own overexpression additionally. This might explain the interesting differences between the transgenic cells' reactions at low and high salt concentrations and, accordingly, their relative performance compared to the wild type.

With reports about possible crosstalks of *pr-10a* with hormone-signaling pathways, for example coincidences of increased abscisic acid (ABA) levels and *pr-10a* expression induction (Genoud and Métraux 1999; Bahramnejad et al. 2010), the regulatory proteins of the Pti class are suspected to be linkers translating the signal from altered plant-hormone levels into modulated PR-gene expression and, thus, triggering physiological changes for stress response (Gu et al. 2002). Further venture in this field should be undertaken for solving the question whether and at which point PR genes interact with plant hormones and which

interactions exist between PR proteins and other connections to hormonal signalling in the cell (Liu and Ekramoddoulah 2006).

Bridging the gap between molecular and phenotypic data

Although a similar expression pattern of *pr-10a* indicated a common stimulus by either applied compound, NaCl as well as sorbitol, the combination of phenotypic information from growth curves and respiration measurements permitted deeper insights into the interaction between the molecular stress response and its impact on superordinate levels. The monitoring of respiration over 24h demonstrated that distinct challenges cause a similar enhancement of respiration, which does, however, not necessarily result in similar growth results. Together with information about the characteristics of the applied chemical compounds, systematic insights into the cells' metabolic activities were possible.

Monitoring the cells' respiration after the long-term pre-treatment in combination with accordingly recorded growth-curve data indicated that enhanced *pr-10a* expression is not beneficial under all circumstances and at all time points. For instance, although the 0.32mM NaCl condition completely impeded growth in all genotypes, the monitoring revealed sustained respiration capacities only in the wild-type cells, whereas both transgenic cell lines exhibited a relatively fast decline. Visual inspection of the cells revealed a higher rate of undyed cell conglomerates in the transgenic lines (data not shown), indicating a higher death rate than in the wild type line.

The thesis clearly demonstrated that respiration measurements are a meaningful complement for the set of methods when examining phenotypes of heterotrophic dedifferentiated cell cultures. The novel method provides valuable information in situations not targeted by more common methods such as growth measurements. Examples in which the new approach is superior include the short-term measurement over 24h, and situations in which the cells' reaction is not measurable as, e.g., under severe stress, where growth is impeded entirely.

Especially the gained information about the very first phase of cellular stress response, which are unattainable by the relative slow growth-measurement approach, provides valuable information about the overall impact of the applied stresses, and the influence of the genetic engineering on it. A major issue is that these experimental results are interpretable only together with information on cell physiology and on the biochemistry of the applied substances.

That is, when working with undifferentiated cells, the specific features of the applied compounds affect the cells directly. In this thesis, this apparent advantage of the chosen experimental set-up resulted in interesting findings regarding altered respiration rates. It would be interesting in future studies to apply polyethylen glycol (PEG) as stressing compound, especially for evaluating responses to osmotic challenges. This non-ionic effector does neither leak into cells nor provide additional energy to the cells (Hohl and Schopfer 1991); the according results would be complimentary to the outcome of the studies conducted in the course of this thesis.

Methodological advances

Vector architecture and choice of suitable promoter

Because applications of dicistronic constructs in plants were few, our knowledge about the impact of the vector architecture and the resulting physiological characteristics of the host-cell system was limited. In the second chapter of this thesis I analysed the effects of several combinations of different 1st cistron elements with different IRES elements on both the translation of the reporter gene and the observable luciferase activity. As major result, it was revealed that the specific combination of 1st cistron element and IRES element determines the success of the 2nd cistron target gene expression. The results were comparable to phenomena of IRES regulation described in Mammalian cells (Hennecke et al. 2001). Thus, I was able to formulate recommendations for devising testing steps and possible solutions for cases were the first cistron gene negatively effects the translational efficiency of the IRES element.

As described above, another feature of the vector system used by El-Banna et al. (2010) for realizing the *pr-10a* overexpression turned out to be problematic for testing salt-stress challenges in the course of this thesis. Namely, the used dicistronic expression cassette is under control of the mannopine-synthase promoter (p-MAS), for which possible induction effects by salt treatments were reported (Langridge et al. 1989; Fox et al. 1992) and confirmed here. As a consequence, results from experiments under salt challenge have to be interpreted with care regarding the role of *pr-10a* in the proposed cause-and-effect chains. To facilitate the measurement of *pr-10a*-overexpression effects in terms of salt challenge unbiasedly, other non-inducible promoters should be chosen for re-designed vector construction in future studies.

Method development for respiratory phenotyping of plant cells

The method development in this thesis for facilitating the surveillance of cell respiration augmented the so far limited repertoire for recording cell cultures' phenotypes by a novel approach particularly suitable for targeting short-term cell responses. The successful application of mitochondrial respiration measurements based on the reduction of a tetrazolium dye (Bochner and Savageau 1977) demonstrated that respiration measurements are applicable to plant cells in analogy to microbiological frameworks based on the OmniLog® PM technique (Bochner et al. 2001, Bochner 2009). In contrast to the traditional recording of growth curves (which is comparatively slow), this approach to measuring provides information about the very first phase of a stress response.

Nevertheless, several critical points concerning the technical realization should be addressed in future research projects. One major point is to ensure the non-toxicity of the dye for the investigated cells. Although in the course of my own studies toxicity was not observed, for other tetrazolium chloride (TTC) derivatives intoxications would not be exceptional (Berridge et al. 2005). Furthermore, it should be taken into account that the dye has to come into contact with the respiration chain in the mitochondrial inter-membrane space. Distinct TTC derivatives may exhibit distinct properties regarding their transportability over membranes. Membrane integrity itself should also be considered, since severe stress may alter the general access of dyes to the cells and the respiration chains.

Additionally, an underlying assumption for the respiration measurements is that electron flow through the mitochondrial respiratory electron chain, which causes the formazan formation, automatically leads to ATP synthesis (Bochner et al. 2001, Bochner 2009). Since the plant mitochondrial respiratory chain can be by-passed using the non-energy conserving alternative oxidase pathway (reviewed in Millar et al. 2011), it was hypothesized that the formed purple colour in TTC based viability assays need not automatically correspond to ATP production rates. Experimental findings concerning the correlation of ATP production and formazan based respiration assays are indeed ambiguous (Duncan and Widholm 2004; Galle et al. 2010; Kurzbaum et al. 2010), thus more research on this topic is necessary.

The technical implementation of the measurement in plant cells is currently inconvenient. The highly automated OmniLog® PM technique (Bochner et al. 2011) for the measurement of microbial respiration makes use of cell suspensions and is processed in 100µl volumes on 96er multi-well plates. In the plant-cell system, the purple-coloured formazan accumulates inside the cells, which makes the respiration measurement over time cost- and labour-

intensive and thus does not yet facilitate a high-throughput approach in analogy to the OmniLog® PM technique. Further research should aim at the discovery of tetrazolium formulations which can escape the cells and thus simplify the process of measurement by superseding the ethanol-extraction step. Also, an adaptation to multi-well plates should be considered. However, the typical plant cell size would favour 24-well plates over 96-well plates.

The here investigated cells were heterotrophic, which makes them independent from energy supply normally emerging from the photosynthesis. Since altered photosynthesis rates and thus perturbations in energy supply are well documented phenomena in the literature on stress research (Gaspar et al. 2002), the transfer of this technique to photoautotrophic cells has to be investigated carefully. For instance, the technical aspect of the absorption optimum being located at 520nm has to be considered as potentially problematic. A bias due to the photometric measurement has to be expected from photosynthetic pigments, but other chromatic cellular compounds have to be taken into account as well.

Coming to terms with longitudinal data characteristics and increasing outcome complexity

With the method developments for plant cell respiration measurements presented in this thesis and regarding the upcoming technical improvements for them, it was conceivable that more sophisticated data-analysis strategies will be soon in need. The last part of this thesis therefore focussed on suitable analysis strategies for data measured over time, i.e. longitudinal data sets of high dimensions. Although based on the output of the OmniLog® PM technique, the tools introduced here are applicable to any type of longitudinal data, irrespective of their origin.

To summarize, a comprehensive and adaptable framework for data visualization and parameter estimation, including 95% confidence intervals for the estimated parameters, was provided. The comparison of the methods for curve fitting (calculating both point estimates and confidence intervals) revealed that the spline approach was more robust against irregular curve shapes than fitting any of the considered models or using the native PM software. Thus, the spline-fitting approach was chosen as a starting point for the automated post-processing of PM data. It was shown that this strategy provides much more information than the strict dichotomization into positive and negative reactions usually applied for PM data analysis.

The results in this thesis form the basis for a forthcoming freely available R package for the analysis of PM data, representative for respiration measurements resulting in longitudinal data

sets. The availability of convenient and robust data exploration techniques via freely available software such as R will allow users to conduct in-depth data analyses that go significantly beyond the consideration of mere endpoint measurements and presence/absence calls.

After selecting and rearranging the data in a sensible way depending on the respective scientific questions of interest, the tools presented in this thesis allow the visualization of the kinetics in several meaningful respects. By further enabling users to autonomously gain the optimal access to the biologically valuable information comprised in respiration kinetics, they are free to set up novel and innovative experimental designs for comparing the cells' respiration behaviour regarding the impact of distinct pretreatments, substrates, genotypes (e.g., wild types vs. genetically engineered lines), and so on. Once the devised convenient graphical representation and parameter estimation features for statistically analysing the resulting data are implemented and published as the forthcoming "opm" R package, a fast and easily applicable tool for analysing high-throughput phenotypic data is available.

Literature

Bahramnejad B, Goodwin PH, Zhang J, Atnaseo C, Erickson LR (2010) A comparison of two class 10 pathogenesis-related genes from alfalfa and their activation by multiple stresses and stress-related signaling molecules. Plant Cell Rep 29: 1235-1250

Berridge MV, Herst PM, Tan AS (2005) Tetrazolium dyes as tools in cell biology: New insights into their cellular reduction. Biotechnology Annual Review 11: 127-152

Bochner BR and Savageau MA (1977) Generalized indicator plate for genetic, metabolic, and taxonomic studies with microorganisms. Applied and Environmental Microbiology 33: 434-444.

Bochner BR, Gadzinski P, Panomitros E (2001) Phenotype MicroArrays for high throughput phenotypic testing and assay of gene function. Genome Research 11: 1246-1255

Bochner BR (2009) Global phenotypic characterization of bacteria. FEMS Microbiol Rev 33: 191-205

Bochner BR, Siri M, Huang RH, Noble S, Lei XH et al. (2011) Assay of the multiple energy-producing pathways of mammalian cells. PLoS ONE 6: e18147

Duncan DR, Widholm JM (1990) Measurement of viability suitable for plant tissue cultures. In: Polland JW, Walker JM (ed.) Methods in Molecular Biology Vol. 6 Plant Cell and Tissue Culture. The Humana Press, New Jersey

El-Banna A, Hajirezaei MR, Wissing J, Zahid A, Vaas L et al. (2010) Over-expression of PR-10a leads to increased salt and osmotic tolerance in potato cell cultures. Journal of Biotechnology 150: 277-287

Fox PC, Vasil V, Vasil IK, Gurley WB (1992) Multiple ocs-like elements required for efficient transcription of the mannopine synthase gene of T-DNA in maize protoplasts. Plant Mololecular Biology 20, 219-233

Galle A, Florez-Sarasal I, Thameur A, de Peape R, Flexas J et al. (2010) Effects of drought stress and subsequent rewatering on photosynthetic and respiratory pathways in *Nicotiana sylvestris* wild type and the mitochondrial complex I-deficient CMSII mutant. Journal of Experimental Botany 61: 765-775

Gaspar T, Franck T, Bisbis B, Kevers C, Jouve L et al. (2002) Concepts in plant stress physiology. Application to plant tissue cultures. Plant Growth Regulation 37: 263-285

Genoud T, Metraux JP (1999) Crosstalk in plant cell signaling: structure and function of the genetic network. Trends in Plant Science 4: 503-507

Gonzales-Lamothe R, Boyle P, Dulude A, Roy V, Lezin-Doumbou C et al. (2008) The transcriptional activator Pti4 is required for the recruitment of a repressosome nucleated by repressor SEBF at the potato *pr-10a* gene. The Plant Cell 20: 3136-3147

Gu YQ, Yang C, Thara VK, Zhou J, Martin GB (2000) Pti4 is induced by ethylene and salicylic acid, and its product is phosphorylated by the Pto kinase. Plant Cell 12: 771-786

Hennecke M, Kwissa M, Metzger K, Oumard A, Kröger A et al. (2001) Composition and arrangement of genes define the strength of IRES-driven translation in bicistronic mRNAs. Nucleic Acids Research 29: 3327-3334

Hohl M, Schopfer P (1991) Water relations of growing maize coleoptiles. Comparison between mannitol and polyethylene glycol 6000 as external osmotica for adjusting turgor pressure. Plant Physiology 95: 716-722

Kurzbaum E, Kirzhner F, Armon R (2010) A simple method for dehydrogenase activity visualization of intact plant roots grown in soilless cultures using tetrazolium violet. Plant Root 4: 12-16

Langridge WH, Fitzgerald KJ, Koncz C, Schell J, Szalay AA (1989) Dual promoter of *Agrobacterium tumefaciens* mannopine synthase genes is regulated by plant growth hormones. Proceedings of the National Academy of Science USA 86, 3219-3223

Liu JJ, Ekramoddoullah AKM (2006) The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses. Physiological and Molecular Plant Pathology 68: 3-13

Millar AH,Whelan J, Soole KL, Day DA (2011) Organization and regulation of mitochondrial respiration in plants. Annual Review of Plant Biology 62: 79-104